Does Size Matter? Bailouts with Large and Small Banks

Eduardo Dávila
Harvard University
edavila@fas.harvard.edu

First Draft: April 28, 2010
This Draft: July 29, 2011

Abstract

Yes, size actually matters. This paper models the strategic interaction between banks and the government when bailouts are possible. I analyze how imperfect common knowledge about the government’s bailout policy affects the ex-ante leverage choice for each bank. Large banks, by internalizing the effect of their size on the likelihood of being bailed out, are willing to take more leverage. In equilibrium, the existence of these large entities induces small banks to also increase their leverage. Therefore, the probability of bailout and the economy-wide leverage is larger when large banks are present. Regulators should treat large banks differently.

JEL numbers: E61, G01, G21, G28

Keywords: bailouts, too big to fail, too many to fail, systemic crisis, bank leverage, banking regulation, global games

*Harvard University, Department of Economics, Littauer Building, 1875 Cambridge Street, Cambridge MA 02138, edavila@fas.harvard.edu. I would like to thank Philippe Aghion, Marios Angeletos, Robert Barro, John Campbell, Emmanuel Farhi, Xavier Freixas, Drew Fudenberg, David Laibson, Ken Rogoff, Andrei Shleifer, David Scharfstein, Alp Simsek, Jeremy Stein, Jean Tirole and Luis Viceira for very helpful comments, as well as participants in Harvard Macro lunch and Harvard/HBS Finance lunch. Financial support from Rafael del Pino Foundation is gratefully acknowledged. Remaining errors are my own.
Motivational quotes

Breaking up big banks wouldn’t really solve our problems, because it’s perfectly possible to have a financial crisis that mainly takes the form of a run on smaller institutions. (...) The next bailout wouldn’t be concentrated on a few big companies - but it would be a bailout all the same.


(...) A surprising number of pundits seem to think that if one could only break up the big banks, governments would be far more resilient to bailouts, and the whole “moral hazard” problem would be muted. That logic is dubious, (...). A systemic crisis that simultaneously hits a large number of medium-sized banks would put just as much pressure on governments to bail out the system as would a crisis that hits a couple of large banks.

Kenneth Rogoff. All for One Tax and One Tax for All? Project Syndicate, 2010-04-29

(...) most observers believe that dealing with the simultaneous failure of many small institutions would actually generate more need for bailouts and reliance on taxpayers than the current economic environment.

Lawrence Summers. Interview with Jeffrey Brown, PBS Newshour, 2010-04-22

1 Introduction

The recent financial crisis has generated heated discussions about the role played by large institutions in financial markets. Many financial agents have been considered “too big to fail”, and while this has led to substantial debate in the public sphere, the number of formal contributions to this topic remains small. This paper shows theoretically how the ex-ante decision of a bank to take on more or less leverage, and accordingly the likelihood of bailout in states of distress, depends crucially on the size distribution of the banking sector through its influence in strategic interactions.

When banks make their funding and borrowing decisions, there is a strong coordination motive: when a systemic shock that prevents refinancing happens, the likelihood of bailout by the government/central bank will be increasing in the amount of outstanding short term
This fact links in equilibrium the return of each bank to the aggregate amount of leverage, creating strong strategic complementarities. A large bank that makes an individual choice is able to get around this inherent coordination problem, since its choice is equivalent to the synchronized decision of many small banks. This fact will push a large bank towards more aggressive positions when it expects a weak policy response. Furthermore, a large bank is aware that its actions may turn out to be pivotal for the possibility of bailout and thus always has an incentive to increase its leverage. When small banks are aware of the presence of these large entities, as long as they reap any benefit from a systemic bailout, they will also be more aggressive in their leverage choice in equilibrium. Overall, the expected leverage for an economy with large banks will be larger than in an equivalent economy with small banks.

With this argument, I have implicitly assumed that large banks are privately better off taking more leverage. This is reasonable, since more leverage increases both profits in good states and the likelihood of bailout in situations of distress. In the model, we can see how individual profitability, information precision about policy and the state of the economy modulate these effects. Even though I provide a simple microfoundation, I have kept the model at a high level of generality, consistent with different explanations for optimal bank capital structure and sources of financial distress.

This paper does not take any stand on why the central bank intervenes in the market and bails out banks, but simply studies the positive implications of an ex-post bailout. In fact, under the assumption that the amount of leverage chosen in a decentralized fashion is inefficient, full commitment not to bailout banks is the ex-ante optimal policy implied by the model. Nonetheless, as long as the central bank effectively bails out the financial institutions in certain circumstances, perhaps because of commitment problems, the mechanism of this paper will be in operation.

I should remark that the often mentioned argument about policymakers giving implicit guarantees to large banks but not small ones is irrelevant for the main mechanism of this paper. That possibility, which can be easily studied inside my model, will simply exacerbate the importance of large players, strengthening my conclusions. By assuming that leverage decisions are made simultaneously, I deliberately abstract from herding and signaling effects, which should also reinforce my conclusions. Moreover, if we think that large banks are better

---

1This paper operates through a simple leverage choice mechanism. The amount of debt outstanding subject to refinancing considerations is assumed to be the right proxy for systemic exposure. A more general formulation in which banks could choose their net exposure in different states would be more realistic, but a similar mechanism to the one discussed in the paper would apply.
informed about central bank policies, they will also exert greater influence in those states where the central banks is thought to be weak, but these considerations are also unnecessary to support the main thesis of this paper.

A simple example can illuminate the underlying mechanism in this paper. Is the government’s decision problem different when it is faced with a bailout of 10 banks of size one versus a bailout of one bank of size 10? If we assume that the risk banks can generate is proportional to their size, the naive answer to this question, from an ex-post perspective, is no. This can be called the “too many to fail” critique to the “too big to fail” problem or the “clones problem” and endorses the view from my three motivational quotes. The problem with this logic is that large banks are aware that their individual choice directly affects the likelihood of bailout, while small banks are individually unable to determine the probability of bailout. Therefore, anticipating the policy response and internalizing the effect of its size, large banks decide to be more aggressive at an ex-ante stage, increasing their leverage in equilibrium and the likelihood of bailout.

To understand the relevant strategic interactions, I build a parsimonious model of the relationship between different financial institutions and the central bank/government. Systemic bailout environments are characterized by strategic complementarities and, as discussed by Cooper and John (1988), this feature can generate multiple equilibria, which makes these problems hard to analyze. I deliberately choose a modeling framework that relies heavily on imperfect information. This choice at once captures a realistic feature of this economic environment and generates unique equilibrium predictions.

This paper is related to the work in global games pioneered by Carlsson and van Damme (1993) and introduced later in the finance and macroeconomics literature by Morris and Shin (1998) and Frankel and Pauzner (2000). A clear exposition of the global game model can be found in Morris and Shin (2003). Important references in banking theory using related methods are Rochet and Vives (2004) and Goldstein and Pauzner (2005): both papers analyze bank runs and depositors decisions, unlike this paper, which focuses on funding decisions by banks and ex-post government bailouts. Angeletos and Pavan (2007) make use of linear quadratic economies to analyze welfare in a related environment; their paper assumes mild complementarities that avoid multiplicity of equilibria and abstract from size considerations. He and Xiong (2010) is a recent contribution that uses similar techniques to analyze debt runs. The textbooks by Vives (2008) and Veldkamp (2009) discuss related topics in environments with imperfect information.

From a technical viewpoint, my results are closely related to previous work by Corsetti
et al. (2004) and Corsetti, Pesenti and Roubini (2002), which analyze the case of currency attacks in the presence of a large speculator, trying to understand the effect caused by George Soros in the Asian crisis of 1997. Their model has a binary action space and only allows for a single large agent. The main technical contribution of this paper is that I solve for any possible combination of large and small agents.

The use of noise in this model should be understood as a relevant feature of the economic environment and not as a perturbation argument to select a particular equilibrium in a given game. I make use of Blackwell’s sufficient conditions and the contraction mapping theorem to provide sufficient conditions on the variance of the noise of the private signals that guarantee a unique equilibrium. These tools, although used to prove existence and uniqueness of equilibrium in other environments, hadn’t been used, to the best of my knowledge, in this literature. This could be a fruitful approach for future applications.

The most recent papers on bailouts, Diamond and Rajan (2009), Farhi and Tirole (2010) and Chari and Kehoe (2010), do not make any conclusive statement about the importance of bank size when bailouts are possible. The older literature on bailouts, for instance, Freixas (1999) or Schneider and Tornell (2004), does not have either strong predictions about how the size of players matters for the outcome. The most closely related paper in the banking literature is Acharya and Yorulmazer (2007), which analyzes the “too many too fail” problem. Although their model is formally very different, some of the main insights highlighted there are captured by my formulation. Neither Freixas and Rochet (2008) nor Gorton and Winton (2003), who offer a detailed textbook treatment of many related topics, present strong theoretical predictions about bank size.

The main contribution of this paper is to recognize that, by internalizing its size, a large financial institution is willing to take more aggressive positions than a small one due to its effect on the likelihood of bailout. Two different channels for large banks arise in this situation: a) the fact that large banks can be pivotal inducing a bailout gives them incentives to always take additional leverage; b) large banks are aware that their individual decisions directly determine the equilibrium bailout probability, while for a small bank the equilibrium bailout probability is set by the equilibrium actions of the other banks and never by its own leverage choices. In equilibrium, the existence of these large financial institutions also modifies the capital structure choice of small ones, tilting them towards taking more aggressive positions.

---

2See the discussion about these issues by Frankel, Morris and Pauzner (2003) in the context of general global games.

3See some examples, for instance, in Vives (2001).
and thus generating more systemic risk, since they feel shielded by the presence of aggressive large banks. Both effects increase the expected amount of leverage and the likelihood of bailouts in a systemic crisis.

Despite the fact that the decision making process of the banks is modeled as extremely rational, the paper can be reinterpreted by assuming that bank management is myopic and that funding markets are the ultimate decision makers over leverage in banks balance sheets. Furthermore, I also suggest how the model could be reinterpreted in terms of lobbying activity.

Section 2 sets up the model and justify its assumptions. Section 3 defines and characterizes the equilibrium for different benchmarks and the general case. Section 4 solves the model numerically, illustrating the main mechanisms of the paper. I discuss some possible extensions of model, its empirical relevance and policy implications in section 5. Section 6 concludes. Appendix A contains technical results, while the online appendix presents several extensions and additional technical material.

## 2 The Model

### 2.1 Banks

The model is a partial equilibrium representation of the banking sector. Only banks, that can be large or small, and the central bank are explicitly modeled. The total amount of equity in the banking sector as a whole is given exogenously and is normalized to be a continuum of measure 1. Small banks, indexed by $k$ and represented as part of a continuum, hold a measure $\int dk = 1 - \lambda$ of the total amount of equity. Large banks, indexed by $j$, are assumed to have noninfinitesimal mass and hold the rest of the equity of the banking sector $\int dj = \lambda$. I denote the set of small banks by $K$ and the set of large banks by $J$. In order to simplify notation, I denote a generic bank, large or small, by $i$ and define $I = K \cup J$. Lastly, I assume that there are $N$ symmetric large banks in $K$, each of them with measure $\frac{\lambda}{N}$. This last assumption about symmetry only simplifies the exposition, but heterogeneity in the size of large banks can be easily incorporated into the model. Figure 1 shows the size distribution of banks graphically.

Banks are risk neutral expected utility maximizers and have a single choice variable $L_i$.

---

4In order simplify the exposition, I make use of the terms bank and central bank throughout the paper, but the model can be applied to all the participants in the financial sector as well as any branch of the government with authority to help financial institutions in situations of distress. Since my focus is on the amount of outstanding short term debt, any institution that relies on short term funding and is taken into account by the policymaker should be considered.
which represents total leverage\textsuperscript{5}. By leverage, I broadly refer to short-term funding with rollover risk and subject to systemic distress. For example, a bank with equity of 5 million and \( L_i = 2 \), is able to issue loans with a value of 10 million funded by 5 million of short term debt in addition to its equity; a bank with \( L_i = 1 \) operates without leverage. Empirically reasonable values for \( L_i \) are around 10 for commercial banks and 20 for investment banks\textsuperscript{6}; see CGFS (2009) for more details. Throughout most of the paper, I assume that bank management, acting on behalf of shareholders, is the ultimate decision maker in each bank. Therefore, each bank’s goal is to maximize its return on equity.

The decision tree that represents the profits per unit of equity for each bank is depicted in figure 2.

All banks choose their leverage \( L_i \) simultaneously at a initial stage. Each unit of short term debt taken faces a fixed repayment at a gross rate of \( 1 + r_i^K \). After each bank decides on \( L_i \), with an exogenously determined probability \( 1 - p \), there is no systemic shock and each loan

\textsuperscript{5}Banking regulation tends to focus on leverage ratios, defined as the the ratio between Tier 1 capital (common shares and retained earnings) and total assets. That definition can be roughly mapped to the inverse of \( L_i \) in my model.

\textsuperscript{6}Depending on the treatment given to deposits and long-term debt, these numbers may vary slightly.
granted generates a gross rate of return $1 + r^L_i$. With probability $p$, a systemic shock hits the banking sector, forcing the central bank to decide whether to bailout the banks or not. $q(\theta, \bar{L})$ denotes the probability of bailout conditional on a systemic shock and its determination in equilibrium is discussed below. If the central bank decides to bailout the banking sector, each bank receives only a fraction $\delta_i$ of the net return on its loans $r^L_i$ and, if there is no bailout, $\gamma_i$ is the fraction of the net return recovered. $\delta_i$ can be thought of as how generous the bailout is and $1 - \gamma_i$ as how important the losses are after the systemic shock hits; I naturally assume that $\delta_i \geq \gamma_i$. Both $\delta_i$ and $\gamma_i$ are exogenous throughout the paper but they could be endogeneized if considered as policy instruments. Whether there is bailout or not, each bank experiences a distress cost of $-\frac{1}{2}\kappa_i L^2_i$ per unit of equity if the systemic shock hits. The distribution of $r^L_i$, $r^K_i$, $\delta_i$, $\gamma_i$ and $\kappa_i$ for the cross section of banks is common knowledge.

A bank with equity $E_i$ is able to grant loans up to $L_i E_i$, therefore, its total profit is given by (1).

$$\pi_i = (1 - p) \left(1 + r^L_i\right) L_i E_i + p \left[ \mathbb{E}_i \left[q(\theta, \bar{L})\right] \left(1 + \delta_i r^L_i\right) L_i E_i - \frac{1}{2}\kappa_i L^2_i E_i \right] + p \left[ (1 - \mathbb{E}_i \left[q(\theta, \bar{L})\right]) \left(1 + \gamma_i r^L_i\right) L_i E_i - \frac{1}{2}\kappa_i L^2_i E_i \right] - \left(1 + r^K_i\right) (L_i - 1) E_i \quad (1)$$

Considerations about the differential cost of capital and investment opportunities between large and small banks can be modeled by varying the cross section of $r^K_i$ and $r^L_i$ in the model. For instance, if we expect small banks to have higher funding costs than large banks, this can be incorporated into the model by assuming $r^K_i > r^L_i$.

Given my assumptions, each bank $i \in I$ maximizes $\pi_i$, the net return on equity accounting for cost of distress, on $L_i$, the amount of leverage. By rearranging (1), the objective function for each bank can be written as in (2).

---

7 I present a simple extension in the online appendix that captures the possibility of vulture behavior: if there is no bailout, less levered banks may benefit from acquiring distressed competitors; in that case, $\gamma_i$ is decreasing in $L_i$.

8 A similar quadratic term could be added to the non-shock state to account in reduced form for risk aversion in bank decisions or other costs unrelated to systemic distress. Moreover, $\kappa_i$ could also be made dependent on experiencing a bailout or not.

9 For small banks, $E_i$ has measure zero: I will only focus on $\frac{\pi_i}{E_i}$, so that is not a concern here.

10 Note that the expectation is taken with respect to the individual information set of each bank. Throughout the paper I use the standard convention $\mathbb{E}_i [\cdot] = \mathbb{E} [\cdot | \text{Information of bank } i]$. Also notice that I discard the term $- (1 + r^K_i) E_i$, since it plays no role in determining $L^*_i$. 

---
\[
\max_{L_i} (1 - p) r^L_i L_i + p \left[ \mathbb{E}_i \left[ q \left( \theta, \bar{L}_i \right) \right] \delta r^L_i L_i + \left( 1 - \mathbb{E}_i \left[ q \left( \theta, \bar{L}_i \right) \right] \right) \gamma_i r^L_i L_i - \frac{\kappa_i L_i^2}{2} \right] - r^K_i L_i \quad (2)
\]

The choice of a linear quadratic environment is deliberate; it is motivated by both technical and economic reasons. From a technical perspective, given the complexity of the equilibrium fixed point, the convenience of a tractable optimal policy rule is extremely important. This formulation delivers a tractable interior solution for \( L_i \) independently of the level of equity of the bank. Note that the cost of distress in (2) depends on the leverage taken by each bank \( L_i \) but not on the total amount of individual debt outstanding. Equivalently, the cost of distress caused by 10 million in outstanding debt is more important for a bank with 2 million in equity \( (L_i = 5) \) than for a bank with 5 million in equity \( (L_i = 2) \). This is a natural assumption. More generally, we can think of the objective function described in (2) as a reduced form approximation of a more complex problem that delivers an interior solution for leverage or, more broadly, for any relevant systemic variable; that view makes the mechanism described in this paper robust to different microfoundations for the choice of capital structure.

I want to justify the assumptions made so far. First, I implicitly assume that long term forces outside the model determine the amount of equity and long term debt banks can raise, mainly due to information asymmetries à la Myers and Majluf (1984), so the relevant decision for the banks in the medium run is to actually determine how much short term debt to take. Analogously, I abstract from directly modeling demand deposits, which has been the focus of an abundant literature in the Diamond and Dybvig (1983) tradition. These shortcuts keep the analysis simple by only identifying short term debt as the single\(^{11}\) systemically relevant variable. When discussing optimal capital regulation for banks, Kashyap, Rajan and Stein (2008) mainly focus on creating mechanisms to reduce the pervasive effects of debt in situations of distress, lending support to this formulation. A more elaborate funding decision should not affect directly the main mechanism presented in this paper.

Second, note that the distribution of equity directly determines the overall size of each bank balance sheet. Nonetheless, if we assume that banking technologies (mainly \( r^L_i \) and \( \kappa_i \) in the model) are roughly homogeneous across banks, leverage ratios will be similar regardless of size. It is also worth noting that, in this model, all the adjustment is carried out through quantities (the amount of leverage) and not through prices or loan returns: there is an implicit

\(^{11}\)In a traditional asset pricing environment with different state contingent claims, banks would have more freedom to choose the correlation of their credit exposures. However, it is always the case that demandable sources of funding can generate correlation of exposures and thus systemic risk.
assumption of a perfectly elastic supply of capital at given rates. In a partial equilibrium context, we can think of each individual $L_i$ as inducing an individual demand function. If we assumed an equilibrium pricing model, different values of $\gamma_i$ and $\delta_i$ would generate cross-sectional dispersion in the funding terms across banks, since they would pay differently in the crisis state.

Third, I consider systemic shocks as independent events from standard business cycles, whose effects are subsumed in $r^L_i$ and $r^K_i$: in a narrow interpretation of the model, systemic shocks are defined as events in which debt refinancing becomes impossible. The fact that $p$ is exogenous and not a function of leverage could be relaxed at the cost of losing tractability, but similar effects to the ones discussed in the paper would apply. If banks were aware of how they can affect $p$, they would also take it into consideration when making their leverage choices.

Last, the losses that occur when a shock hits the banking sector and that account for the quadratic loss term $-\kappa_i L^2_i$, can be rationalized with different mechanisms. Natural explanations can be an increase in adverse selection that leads to market breakdowns, as in Akerlof (1970), Dang, Gorton and Holmstrom (2009) or Tirole (2010) or more recent Knightian uncertainty arguments as in Caballero and Krishnamurthy (2008) or Caballero and Simsek (2009). Furthermore, fundamental or expectations driven bank runs in deposits triggered by the losses induced by the inability of rolling over debt can also be considered as the source of these losses. Explicitly microfounded models about rollover risk and debt and repo runs are presented in recent work by Acharya, Gale and Yorulmazer (2010), Acharya and Viswanathan (2011), He and Xiong (2010) and Martin, Skeie and Von Thadden (2010). Search frictions as those described in Lagos, Rocheteau and Weill (2008) could also be relevant. Any of the private costs analyzed in this literature suffice as an explanation for the loss term. An implicit assumption for any of these explanations is that banks are engaging in some kind of maturity transformation with their short term demandable or hard to renew liabilities, forcing them to liquidate positions or find new sources of funding on unfavorable terms when a shock hits them.

### 2.2 Central Bank Policy and Information Structure

The probability of bailout $q(\theta, L)$ depends on the policy response of the central bank and is determined in equilibrium. Formally, the government follows the policy:
\[ q(\theta, \tilde{L}) = \begin{cases} 
1, & \tilde{L} \geq \theta \\
0, & \tilde{L} < \theta 
\end{cases} \]

where \( \tilde{L} \) is defined as the global amount of leverage in the system \( \tilde{L} = \int L_idi \) and \( \theta \) denotes the type of the central bank. This parameter \( \theta \) modulates the likelihood of being bailed out by the central bank. We can think of \( \theta \) as given by the underlying fundamentals of the economy: an economy has a large \( \theta \) if the the costs of allowing for systemic defaults are small. Alternatively, \( \theta \) can be interpreted as the hawkishness of the central bank, that is, the willingness of the central bank to bailout the financial sector given the state of the economy. This last interpretation is similar to the Diamond and Rajan (2009) household friendly/institution friendly categorization of the central bank.

A policy response like (3) implies that if the amount of leverage in the system is large in situations of distress, the central bank is forced to step in and bail out the whole banking sector: this policy response creates strategic complementarities among banks. Allowing for smoother or more general policy responses does not modify my conclusions, as long as the complementarities remain.

I want to remark that the policy response in (3) fully incorporates the “too many too fail argument”. For the central bank it is completely irrelevant whether a single large bank or many small banks are exposed in a systemic shock: only the total amount of debt outstanding is what actually matters for policy.

An interesting feature of the model is that even if a single bank in the continuum takes an arbitrarily large leveraged position, it will not affect the global amount of outstanding short term debt in the banking sector; this is not the case for large banks. A more general formulation could assume that the relevant amount of leverage is \( \int w_i L_idi \), where \( w_i \) denotes institution specific weights. If, in addition to their size, we expect large banks have a larger weight for the policymaker \( w_i \), they would take this into consideration when making decisions, strengthening my results. Small banks cannot internalize this effect directly, but it would certainly play a role in equilibrium.

Note that \( \int L_idi \) refers to the total size of the outstanding positions and not only to the net positions of the banking sector with the rest of the economy. This is consistent with the view that large and complex financial systems are more likely to be fragile and be on the need of government intervention.

The parameter \( \theta \) is ex-ante random. As shown by Angeletos, Hellwig and Pavan (2006), policy interventions in \( \theta \) in this kind of environments are not trivial due to signaling issues,
but they could be potentially addressed. I discuss policy implications of the model in section 6.

The information structure in the model is standard. The value of $\theta$ is initially drawn from an improper prior\textsuperscript{12}.

$$\theta \sim U [\, -\infty, +\infty \,]$$

Each small bank observes a private signal $x_k$ of the type of the central bank $\theta$ and each large bank observes a private signal $x_j$ with possibly different precision.

$$x_k = \theta + \sigma_k \varepsilon_x$$
$$x_j = \theta + \sigma_j \varepsilon_x$$

The noise term $\varepsilon_x$ follows a standard normal distribution with independent realizations across banks. This distributional assumption is not crucial, but it makes the analysis tractable. Note that, in this model, any subset of small banks with positive measure could potentially recover the true value of $\theta$ by pooling their private signals. Therefore if we think about merging small banks to create a large bank inside the model, we must also assume that some informational friction\textsuperscript{13} prevents the merged large bank from learning the actual value of $\theta$. Later on, especially when performing comparative statics on $\lambda$, we should keep in mind that, in this model, having larger banks entails a reduction in the amount of economy-wide information.

An important question is how can we map a bailout in this model to reality. Following the discussion in Farhi and Tirole (2010), I categorize bailouts as monetary or fiscal bailouts, and this model can encompass both. The most natural interpretation for a bailout inside the model is as a monetary bailout. In that case, the central bank simply decides to lower interest rates below the socially optimal level, helping all the distressed institutions to find cheaper borrowing to continue operating. The popular press has referred to this central bank behavior as the “Greenspan put” or, later, as the “Bernanke put”. This mechanism works symmetrically for all the distressed institutions creating strong complementarities.

The case of a fiscal bailout may seem harder to embed in my framework. In that case, the central bank directly injects equity into the banks, buy assets from them or directly transfers

\textsuperscript{12}It is possible to allow for a more general prior/public signal, but this just makes harder to convey the main intuition in the paper. I briefly discuss the public signal case in the online appendix. As usual, multiple equilibria can be restored when public information is too precise relative to private information.

\textsuperscript{13}The informational frictions described in Stein (2002) could perfectly apply in this context.
resources, but this is an institution specific policy that is not necessarily the same for all banks. There are two ways to understand these policies inside the model: a) the central bank does not discriminate who is going to get public support and implements nontargeted policies, e.g., TALF, TARP or similar programs; b) we can assume that, ex-ante, each institution doesn’t know if it will be bailed out or not, even when it is known that the government will decide to bailout someone\(^{14}\) in a crisis state. The possibility that the central bank allows some banks to fail will be reflected in a lower parameter \(\delta_i\). An important takeaway is that, from an ex-ante perspective, it doesn’t matter which bank exactly gets bailed out, only the average probability. This last interpretation hasn’t been previously discussed in detail in the literature, but it emerges naturally from the model.

The fact that bailouts are nontargeted is essential for my results. However, an adequate cross-sectional distribution for the parameters \(\gamma_i\) and \(\delta_i\) could fit additional plausible bailout policies. For example, it is often argued that large institutions receive implicit “too big too fail” bailout guarantees but small institutions do not. That bailout scheme can be incorporated into the model by assuming that \(\delta_j > \delta_k\), that is, the bailout recovery rate is larger for large banks than for small ones\(^{15}\). That particular bailout specification would induce large banks to take on additional leverage, magnifying all the effects described in the paper.

2.3 Individual optimality and regularity conditions

Before studying the equilibrium of the model, I first discuss the individual problem for a bank. The optimal leverage choice for a bank that observes a signal \(x_i\) is given by the first order condition (4).

\[
L^*_i(x_i) = \left\{ 1 - p + p \left( \gamma_i + (\delta_i - \gamma_i) \left[ \mathbb{E}_i \left[ q(\theta, \bar{L}) \right] \right] + L^*_i \frac{\partial \mathbb{E}_i \left[ q(\theta, \bar{L}) \right]}{\partial L_i} \right) \right\} \left\{ \frac{r^L_i}{p\kappa_i} - \frac{r^K_i}{p\kappa_i} \right\}
\]

This is the key equation of this paper. First, the difference between the return on loans and the cost of capital, \(r^L_i - r^K_i\), controlling for the expected cost of distress \(p\kappa_i\), determines the overall scale of leverage taken. The term \((1 - p)\) accounts for the profits in the no crisis state, while the term \(p \left( \gamma_i + (\delta_i - \gamma_i) \left[ \mathbb{E}_i \left[ q(\theta, \bar{L}) \right] + L^*_i \frac{\partial \mathbb{E}_i \left[ q(\theta, \bar{L}) \right]}{\partial L_i} \right) \right)\) denotes the additional profit

\(^{14}\)Having as example the 2008 crisis, we can think of the Bear Stearns bailout vs. the Lehman collapse.

\(^{15}\)Assuming a larger weight for large banks \(w_j\) for the policy decision would generate similar results.
in the crisis state. This term contains both the individual expectation of bailout given the
information of each bank and the increase in the likelihood of bailout induced by the marginal
unit of leverage. The marginal change in the expected likelihood of bailout
\( \frac{\partial E_i[q(\theta, \tilde{L})]}{\partial L_i} \) is
multiplied by \( L_i^* \), which implies that this effect is stronger for already high levels of leverage.

The comparative statics in exogenous parameters are exactly as expected:
\( \frac{\partial L_i^*}{\partial \delta_i} \geq 0, \frac{\partial L_i^*}{\partial \gamma_i} \geq 0, \frac{\partial L_i^*}{\partial p} \leq 0. \)

Condition (4) applies in general, but small banks, given their infinitesimal size, always face
\( \frac{\partial E_i[q(\theta, \tilde{L})]}{\partial L_i} = 0. \) Conditional on other players actions and its private signal, the small bank problem is convex\(^{16} \), so (4) suffices to characterize the equilibrium response for small banks.

On the contrary, the problem for large banks can potentially be nonconvex and thus requires
additional conditions to be well behaved. First of all, in order for \( \frac{\partial E_i[q(\theta, \tilde{L})]}{\partial L_i} \) to be well defined,
the conditional expectation must be smooth in \( L_i \), so the marginal approach used in (4) is not
valid in the common knowledge case. Introducing any amount of noise about \( \theta \) or about the
equilibrium actions of other players will solve this concern. Therefore, in order to guarantee
the existence of unique global optimum for the large bank individual problem, I make the
following assumption:

**Assumption 1 (enough noise for large banks):** The signal of the large banks is noisy
enough; that is, \( \sigma_j \) is sufficiently large, so that (5) holds.

\[
2 \frac{\partial E_j[q(\theta, \tilde{L})]}{\partial L_j} + L_j^* \frac{\partial^2 E_j[q(\theta, \tilde{L})]}{\partial L_j^2} < \frac{1}{p(\delta_j - \gamma_j)} \frac{\ell_j}{\sigma_j} \quad \forall x_j
\]  

(5)

By guaranteeing that the second order condition (5) holds, we are just making the problem
in (2) globally convex. On intuitive grounds, an arbitrarily well informed large bank will have
to have two local maxima depending on whether it believes that there is going to be a bailout
or not; assuming enough noise smooths out this nonconvexity. The case with a single large
bank (\( \lambda = N = 1 \)), in which \( E_j[q(\theta, \tilde{L})] \) can be expressed analytically, provides clear intuition
about the role played by \( \sigma_j \) not being too small. I present this case in the online appendix.

From a practical viewpoint, given that the amount of noise faced by financial institutions
about policy decisions and about the actions of other market participants in bailout
environments is far from negligible, this assumption can be easily defended empirically.

\(^{16}\)Given a signal, small banks form an expectation about \( q(\theta, \tilde{L}) \), what makes them face a constant marginal
benefit of \( L_i \). This fact and the increasing marginal cost of leverage yields a well defined convex problem.
3 Equilibrium characterization

In this section, I characterize the equilibrium of the model for various benchmarks and then for the most general case. I discuss sufficient conditions for uniqueness in Appendix A.

I have laid out the model under the most general set of assumptions about parameter values, with the goal of stressing how this this framework can fit many different stories about bank behavior without obscuring the mechanism proposed in the paper. From now on, to simplify the exposition, I solve and discuss a stripped down version of the general framework by assuming $\gamma_i = 0$, $\delta_i = 1$, $r^K_i = 0$ and $r^L_i = p\kappa_i$. These assumptions scale the leverage decisions around $L_i \in [0,1]$. None of the robust results of the paper rely on the symmetry of payoffs.

This parameter specification reduces (2) to (6).

$$\max_{L_i} (1 - p) L_i + p E_i \left[ q \left( \theta, \bar{L} \right) \right] L_i - \frac{L_i^2}{2} \quad (6)$$

Under these assumptions, the optimal behavior for small banks collapses to

$$L^*_i = 1 - p + p E_k \left[ q \left( \theta, \bar{L} \right) \right] \quad (7)$$

And for large banks to

$$L^*_j = 1 - p + p \left( \mathbb{E}_j \left[ q \left( \theta, \bar{L}, L^*_j \right) \right] + L^*_j \left. \frac{\partial \mathbb{E}_j \left[ q \left( \theta, \bar{L}, L^*_j \right) \right]}{\partial L_j} \right|_{L_j = L^*_j} \right) \quad (8)$$

I have made explicit in (8) that the probability of bailout depends on the choice $L^*_j$ and that this is taken into consideration by the large bank. The distinction between internalizing size or not is clear when comparing (7) and (8): large banks take into account their effect in the equilibrium value of $q(\cdot)$ through $\mathbb{E}_j \left[ q \left( \theta, \bar{L}, L^*_j \right) \right]$ and additionally they consider how the possibility of being pivotal can modify their returns by $L^*_j \left. \frac{\partial \mathbb{E}_j \left[ q \left( \theta, \bar{L}, L^*_j \right) \right]}{\partial L_j} \right|_{L_j = L^*_j}$. When solving the model numerically, I will refer to the first mechanism as the equilibrium effect and to the second one as the direct effect.

$^17$Small banks choose leverage $L^*_i \in [0,1]$, but large banks may take leverage larger than 1 when they can be pivotal and modify the likelihood of bailout.
3.1 Benchmark: Common knowledge about $\theta$

As a benchmark, it is useful to first solve the model with common knowledge of $\theta$. In the most general case with a continuum of small banks and $N$ large banks at the same time, i.e., $\lambda \in (0, 1)$ and $N \geq 1$, the solution is as follows:

1. When $\theta \leq 1 - p + \frac{p \lambda}{N}$, there is a unique equilibrium in which every bank, large or small, chooses $L_i^* = L_k^* = 1$ and there is always a bailout $q(\theta, \bar{L}) = 1$.

2. When $\theta \geq 1$, there is a unique equilibrium in which every bank, large or small, now chooses $L_i^* = L_k^* = 1 - p$ and $q(\theta, \bar{L}) = 0$.

3. When $1 - p + \frac{p \lambda}{N} < \theta < 1$, there are multiple equilibria. One type of equilibria is characterized by $L_j^* = L_k^* = 1 - p$ and $q(\theta, \bar{L}) = 0$ and the other is given by $L_j^* = L_k^* = 1$ and $q(\theta, \bar{L}) = 1$.

Figure (3) represents the solution graphically.

Uniqueness $\theta \leq 1 - p + p \frac{\lambda}{N}$  Multiplicity $\theta \in (1 - p + p \frac{\lambda}{N}, 1]$  Uniqueness $\theta > 1$

Figure 3: Equilibrium regions with common knowledge

In general, as long as $\theta \leq 1 - p$, $L_i^* = 1$ is a dominant strategy. Under perfect information, it is always the case that $L_i^* \in [1 - p, 1]$, which implies that, when $\theta \geq 1$, the optimal solution is to be as conservative as possible. When there are only small banks, we go back to a classic coordination game where the rest of the parameter space features multiplicity. A large bank by itself can reduce this region by using its mass, with the extreme case given by $\lambda = 1$ and $N = 1$. In that case, the large bank completely eliminates the region of multiplicity. Observe that it is the size of the largest bank that matters to determine the regions of uniqueness and multiplicity with common knowledge.

The measure of the multiplicity region is given by $B = \left| 1 - (1 - p + p \frac{\lambda}{N}) \right| = p \left| 1 - \frac{\lambda}{N} \right|$, the comparative statics are as expected, $\frac{\partial B}{\partial p} \geq 0$, $\frac{\partial B}{\partial \lambda} \leq 0$ and $\frac{\partial B}{\partial N} \geq 0$. This type of solution is not surprising for a coordination game with strong complementarities. The introduction of imperfect information eliminates this multiplicity.
3.2 Only small banks ($\lambda = 0$)

Having analyzed the benchmark solution with common knowledge, I turn to the different cases with imperfect information. First I focus on the case with only a continuum of small banks and no large banks, i.e., $\lambda = 0$.

As described above, each small bank observes a signal $x_k = \theta + \sigma_k \varepsilon_x$, with $\varepsilon_x$ being independent standard normal random variable, and creates beliefs over $\theta$, which is drawn from an improper prior $\theta \sim U[-\infty, +\infty]$. Consequently, from the perspective of bank $k$, $\theta | x_k \sim N(x_k, \sigma_k)$, so that $E_k[q(\theta, \bar{L})] = E_k[1(\theta \leq \bar{L}(\theta))] = \int 1(\theta \leq \bar{L}(\theta)) d\Phi(\theta | x_k) = \int 1(\theta \leq \bar{L}(\theta)) \frac{1}{\sigma_k} \phi \left( \frac{\theta - x_k}{\sigma_k} \right) d\theta$, where I have defined $\bar{L}(\theta) = \int L_k(x_k) d\Phi(x_k(\theta))$ as the aggregate amount of short term debt given $\theta$. The fact that $\bar{L}(\theta)$ is a deterministic function of $\theta$ comes from assuming a law of large numbers, given that small banks form a continuum. This allows us to define the equilibrium as a function of a threshold $\theta^*$.

**Definition 1.** An **equilibrium** is defined as

1. An optimal policy rule $L_k^*(x_k)$ that satisfies, for all $x_k$

   $$L_k^*(x_k) = 1 - p + p \left[ \Phi \left( \frac{\theta^* - x_k}{\sigma_k} \right) \right]$$

2. A threshold $\theta^*$, such that $\theta^* = \bar{L}(\theta^*)$, defined by

   $$\theta^* = 1 - p + p \int_{-\infty}^{+\infty} \Phi \left( \frac{\theta^* - x_k}{\sigma_k} \right) d\Phi \left( \frac{x_k - \theta^*}{\sigma_k} \right) = 1 - p + \frac{p}{2} = \bar{L}^*(\theta^*)$$

where I have used the fact that the left hand side of the last equation is constant in $\theta$. This result and the fact that a law of large numbers can be used to show that $\int L_k^* d\Phi(x_k(\theta)) \to_p \bar{L}^*(\theta)$, are crucial for this equilibrium characterization. In this particular case, when noise is made arbitrarily small $\sigma_k \to 0$, the model still generates a unique equilibrium prediction. Taking that particular limit, the optimal leverage function approximates a step function with $L_k^* = 1$ when $x_k \leq \theta^*$ and $L_k^* = 1 - p$, when $x_k > \theta^*$, a very different outcome than the one under perfect information.

3.3 A single large bank ($\lambda = 1$ and $N = 1$)

Next, I study the benchmark with a single large bank, i.e., $\lambda = 1$ and $N = 1$. This is a single agent decision problem, so the large bank fully internalizes that it is the single decision maker and also the fact that it is able to change the likelihood of being bailed out. Given assumption 1, the problem has a unique solution for each $x_j$. 

16
The large bank observes a signal \( x_j = \theta + \sigma_j \varepsilon_x \) and its optimal choice of leverage is given by (11).

\[
L_j^*(x_j) = 1 - p + p \left( \Phi \left[ \frac{L_j^*(x_j) - x_j}{\sigma_j} \right] + L_j^*(x_j) \frac{1}{\sigma_j} \phi \left[ \frac{L_j^*(x_j) - x_j}{\sigma_j} \right] \right)
\]

(11)

where we have used the fact that \( \mathbb{E}_j [q(\theta, L_j)] = \mathbb{E}_j [1(\theta \leq L_j(x_j))] = \Phi \left[ \frac{L_j(x_j) - x_j}{\sigma_j} \right] \). In order to find the equilibrium policy \( L_j^*(x_j) \), we must solve the fixed point in (11). An intuitive way to understand (11) is by rewriting it in a multiplier formulation

\[
L_j^*(x_j) = \frac{1}{1 - \frac{p}{\sigma_j} \phi \left[ L_j^*(x_j) - x_j \right]} \left( 1 - p + p \Phi \left[ \frac{L_j^*(x_j) - x_j}{\sigma_j} \right] \right)
\]

The "enough noise" assumption guarantees then that the multiplier is well defined. The optimal leverage choice for the single large bank can be seen as that for a small bank modified by a multiplier term that depends on the marginal change in the likelihood of being bailed out, in addition to the fact that \( L_j^*(x_j) \) is directly determined by the large bank, unlike \( \theta^* \).

3.4 N symmetric large banks (\( \lambda = 1 \) and \( N \geq 1 \))

This situation is harder to analyze than the case with a continuum of banks, since we cannot assume a law of large numbers to pin down the global amount of leverage in equilibrium by conditioning on \( \theta \). With \( N \) symmetric large banks, total leverage \( \bar{L} = \frac{1}{N} L_j + \sum_{j' \neq j} \frac{1}{N} L_{j'}(x_{j'}) \). Therefore,

\[
\mathbb{E}_j \left[ q(\theta, \bar{L}) \right] = \mathbb{E}_j \left[ 1\left( \theta \leq \bar{L} \right) \right] = \mathbb{E}_j \left[ 1\left( \theta - \sum_{j' \neq j} \frac{1}{N} L_{j'}(x_{j'}) \leq \frac{1}{N} L_j(x_j) \right) \right] = \mathbb{E}_j \left[ S|x_j \leq \frac{1}{N} L_j(x_j) \right] = F_{S|x_j} \left( \frac{1}{N} L_j(x_j) \right).
\]

I have defined \( S|x_j \) as the random variable \( \theta - \sum_{j' \neq j} \frac{1}{N} L_{j'}(x_{j'}) \) conditional on \( x_j \), with cdf \( F_{S|x_j} \) and pdf \( f_{S|x_j} \). The fact that \( S|x_j \) is not normally distributed, since \( L_{j'}(x_{j'}) \) is not normal, simply makes the computational exercise more difficult, since \( S|x_j \) must be simulated, but changes nothing from an economic perspective.

**Definition 2.** An equilibrium is defined as an optimal policy rule \( L_j^*(x_j) \) that satisfies, for all \( x_j \)

\[
L_j^*(x_j) = 1 - p + p \left( F_{S|x_j} \left( \frac{1}{N} L_j^*(x_j) \right) + L_j^*(x_j) \frac{1}{N} f_{S|x_j} \left( \frac{1}{N} L_j^*(x_j) \right) \right)
\]

When \( N \to \infty \) the model approaches the case with a continuum of small banks. Intuitively, the term corresponding to size converges to zero and by a law of large numbers the equilibrium policies converge to a given function of \( \theta \), as shown above.
3.5 General case ($\lambda \in [0, 1]$ and $N \geq 1$)

After gaining some intuition with particular cases, this section discusses the general case that features both large and small banks. Recall that $i$ indexes a generic bank, $j$ large banks and $k$ small banks. All other cases are special parameterizations of this one.

The global amount of leverage is given by $\bar{L} = \int L_k dk + \frac{\lambda}{N} \sum_{j=1}^{N} L_j$. In equilibrium we must expect that, given $\theta$, $\int L_k dk \to_p (1 - \lambda) L(\theta)$ by a law of large numbers. For the large banks, given a signal $x_j$, the probability of bailout is given by

$$E_j \left[ q \left( \theta, \bar{L} \right) \right] = E_j \left[ 1 \left( \theta \leq \bar{L} \right) \right] = E_j \left[ 1 \left( \theta - (1 - \lambda) L(\theta) - \sum_{j' \neq j} \frac{\lambda}{N} L_{j'}(x_{j'}) \leq \frac{\lambda}{N} L_j(x_j) \right) \right] = E_j \left[ 1 \left( S|x_j \leq \frac{\lambda}{N} L_j(x_j) \right) \right] = F_{S|x_j} \left( \frac{\lambda}{N} L_j(x_j) \right)$$

Analogously to the case with only $N$ large banks, I have defined $S|x_j$ as the random variable $\theta - (1 - \lambda) L(\theta) - \sum_{j' \neq j} \frac{\lambda}{N} L_{j'}(x_{j'})$ conditional on $x_j$, with cdf $F_{S|x_j}$ and pdf $f_{S|x_j}$.

For small banks, given a signal $x_k$, the probability of bailout is given by

$$E_k \left[ q \left( \theta, \bar{L} \right) \right] = E_k \left[ 1 \left( \theta \leq \bar{L} \right) \right] = E_k \left[ 1 \left( \theta - (1 - \lambda) L(\theta) - \sum_{j=1}^{N} \frac{\lambda}{N} L_j(x_j) \leq 0 \right) \right] = E_k \left[ 1 \left( S|x_k \leq 0 \right) \right] = F_{S|x_k} (0)$$

I have defined $S|x_k$ as the random variable $\theta - (1 - \lambda) L(\theta) - \sum_{j=1}^{N} \frac{\lambda}{N} L_j(x_j)$ conditional on $x_k$, with cdf $F_{S|x_k}$ and pdf $f_{S|x_k}$.

Any bank, large or small, observes a signal $x_i$ and then forms joint beliefs over $\theta$ and the leverage choices of the other banks. This implies that a small bank must form beliefs over $N$ leverage choices in addition to the value of $\theta$ and a large bank must do the same but only with $N - 1$ leverage choices. When forming conditional beliefs, given a realization for $x_1$, $\theta|x_i \sim N(x_i, \sigma_i^2)$ and $x_{i'}|x_i \sim N(x_i, \sigma_i^2 + \sigma_{i'}^2)$ are jointly normally distributed$^{18}$, with a correlation of $\sigma_i^2$ between $\theta|x_i$ and $x_{i'}|x_i$. Given the optimal policy rules, everything must be consistent in equilibrium.

**Definition 3.** An equilibrium is defined by

1. A policy rule $L_k^* (x_k)$ for the small bank.

\[
L_k^* (x_k) = 1 - p + pE_k \left[ 1 \left( S|x_k \leq 0 \right) \right] = 1 - p + pF_{S|x_k} (0) \tag{12}
\]

---

$^{18}$See more details in the online appendix.
2. A policy rule $L^*_j(x_j)$ for the large bank.

$$L^*_j(x_j) = 1 - p + p \left( F_{S|x_j} \left( \frac{\lambda}{N} L^*_j(x_j) \right) + L^*_j(x_j) \frac{\lambda}{F_{S|x_j}} \right)$$

(13)

4 Numerical results

Despite the simplicity of the model, as we move away from the case with only a continuum of agents, finding an equilibrium involves a complicated functional fixed point, so it is not possible to find analytical solutions. I have to rely on numerical techniques to solve for the equilibrium functions. The online appendix describes the numerical procedures in detail.

4.1 Understanding the mechanism: decomposing optimal leverage

Before showing the numerical solutions for the general case, it seems natural to understand the basic mechanisms that induce the different equilibrium policies for large and small banks. Figure 4 plots the optimal leverage function in three different cases. The first is the optimal leverage choice for a small bank when there is only a continuum of small banks in the economy, i.e., $\lambda = 0$, so equation (9) is the relevant one; the second is the optimal leverage choice for a single large bank that fully optimizes, i.e., $\lambda = 1$ and $N = 1$, represented by equation (11) and the last is the optimal leverage choice for a single large bank that behaves myopically, this is, a large bank which behaves according to equation (11) but sets the direct effect term $L^*_j(x_j) \frac{1}{\sigma_j} \phi \left[ \frac{L^*_j(x_j) - x_j}{\sigma_j} \right]$ to zero. Understanding (9) and (11) is essential to interpret figure 4.

What exactly is a large bank in this model and how is it different from many small banks? On the one hand, a large bank acts as a set of small banks that receive perfectly correlated signals. On the other hand, a large bank fully internalizes that its actions have the same weight in the policy as a coordinated set of small banks of the same size. The comparison of the three optimal policies in figure 4 allows to disentangle these two effects.

Let’s first compare the behavior of the small banks with the myopic large bank, which can be thought of a set of small banks with perfectly correlated signals; this eliminates the direct effect term. Small banks know that the aggregate probability of bailout is determined by $\theta^*$, an equilibrium value that they take as given; on the contrary, a single large bank is aware

\[19\text{A brief comment about techniques here: I solve the system of functional equations by using a collocation method with cubic splines, relying on Gaussian quadrature when possible. Unfortunately, for the general case, only Monte Carlo methods are feasible, what considerably slows down the procedure. I solve the nonlinear system of equations with Broyden’s method.}\]
that its leverage \( L^*_j(x_j) \) is the only determinant of aggregate leverage and, in consequence, of the actual probability of bailout. This makes the large bank more aggressive for low values of the signal and more prudent for high values. Intuitively, a single large bank acknowledges that when it is being aggressive the likelihood of bailout is increased by its action and vice versa when it is not. A small bank knows that its particular leverage can never modify the equilibrium probability of bailout directly.

How does a large bank that behaves optimally differ from a myopic one? I have described the difference as the direct effect, which is never present for small banks. This direct effect accounts for the fact that a marginal unit of leverage may discretely induce a bailout and it always makes large banks more aggressive. Note that this effect works on top of the equilibrium considerations described in the previous paragraph.

Furthermore, this direct effect is the source of an interesting nonmonotonicity of the optimal amount of leverage in the signal \( x_j \). The intuition for this result is the following: when it receives a signal that indicates that the central bank is hawkish, a large bank may believe that it is in the region where it’s behavior can be pivotal for a regime switch; in that case, an additional unit of leverage may carry the extra benefit of making all the inframarginal leverage units profitable if a bailout succeeds. Note that this effect is only relevant in the region where a bailout is plausible and it is not present when the government is thought to be either too weak or too strong. This nonmonotonicity, absent in any previous work, is a caused by my assumptions about profits, specifically particular, the possibility of local nonconvexities in the
large bank objective function. When the importance of large banks is small, that is, when \( \lambda \) is small or \( N \) is large, the nonmonotonicity disappears but the direct effect still plays an important role to make large banks more aggressive.

In summary, being large has two different implications: first, a large bank is equivalent to many small banks that are aware of sharing the same signal, directly recognizing the effect of their actions in the equilibrium probability of bailout and getting around the coordination problem; second, a large bank acknowledges the possibility of being pivotal in inducing a bailout. The interaction of these effects determines the difference in behavior between large and small banks.

To ease the exposition, I will focus hereafter only on the equilibrium that arises with fully optimizing large banks, in which all the effects act simultaneously.

### 4.2 Symmetrically informed banks \( \sigma_j = \sigma_k \)

I choose \( p = 1 \) and \( \lambda = 0.75 \) as the reference values for the simulations. This choice has been made to highlight the intuition behind the model and to present intuitive figures, rather than to match banking industry actual data. Varying \( p \) does not change the nature of the problem; it simply scales up or down the effect of the bailout state. \( \lambda \) always enters as \( \frac{\lambda}{N} \) and a reasonably large value for this ratio is needed if we want large banks to play an important role. I first present the results with equally informed banks \( \sigma_j = \sigma_k = 1 \) and then discuss the case in which large banks have more precise information. The online appendix presents alternative calibrations.

#### 4.2.1 Optimal leverage choice and expected equilibrium leverage

Figure 5 shows the optimal leverage choices in equilibrium, \( L_k^* \) and \( L_j^* \), given a private signal \( x_k \) or \( x_j \) for the case with 2 large banks, \( N = 2 \) and \( \lambda = 0.75 \). I plot as benchmarks the cases with only a continuum of small banks, \( \lambda = 0 \), and the one with a single large bank, \( \lambda = 1 \) and \( N = 1 \).

From the large banks’ perspective, having small banks in the economy flattens their optimal leverage choice, making them less aggressive for low values of \( x_j \) but more aggressive when their signal is high. This is because the large banks’ own actions are less relevant in the determination of equilibrium leverage in the presence of small banks. For example, when the large banks obtain signals that indicate a weak policymaker, they have to take into account that small banks may think the opposite and vice versa when the signal indicates a hawkish
policymaker. On the contrary, the presence of large banks makes small banks more aggressive for any realization of their private signal. Intuitively, small banks are aware that large banks, which tend to be more aggressive than small banks for a given signal, are present. Since they are uncertain about both $\theta$ and $x_j$, small banks feel shielded by the large banks’ aggressive policies and have an incentive to also be more aggressive. Noise is crucial for this result: although for a small region of private signals large banks are more prudent, small banks are uncertain about the signal received by large banks, so it is the larger average leverage by large banks what influences small bank behavior.

To better understand the differences between the benchmark situation with only a continuum of banks and the other cases, I plot in figure 6 the difference between the optimal rules. The left plot shows the difference between the optimal leverage for a small bank when large banks are present and the optimal leverage for a small bank when there is only a continuum of banks, $\lambda = 0$. The right plot shows a) the difference between the optimal leverage for a single large bank, $\lambda = 1$ and $N = 1$, and the optimal leverage for a small bank in the continuum and b) the difference between the optimal policy rule for a large bank when there are large and small banks and the optimal leverage for a small bank when there is only a continuum of banks, $\lambda = 0$. These two figures reinforce the intuitions discussed above.

20 The difference between the optimal leverage choices for a small bank in an economy with large and smalls versus an economy with only a continuum of banks is hard to visualize in figure 5.

21 Note that the Y axis in the left plot is an order of magnitude smaller.
An additional variable of interest is the expected amount of leverage in equilibrium for each value of the fundamental $\theta$. Even though the policy response in this model follows a discrete configuration, it may be the case that the social costs of a bailout depend directly on the amount of distress caused by the short term debt refinancing problems and not only on whether there is an ex-post bailout or not. Figure 7 presents the expected amount of leverage for different values of $\theta$. In order to calculate this function, I simply consider each value of $\theta$ and then take an expectation over leverage in equilibrium given the received signals by each bank. Unlike the two previous figures, the horizontal axis now represents $\theta$ and not $x_i$.

As we could have expected from the behavior discussed above, it is always the case that the expected amount of leverage is larger when there are large banks. Even though large banks are less aggressive when they expect a tough central bank, since for most of their signals they are actually more aggressive than small banks and they are always uncertain about other banks signals and the policy response, the expected amount of leverage averages out to a larger figure when large banks are present.

4.2.2 Bailout probability

So far, I have focused exclusively on the individual choices made by each bank. Below, I answer the important question of whether the presence of larger banks increases the probability of bailout in the case of a crisis. Since I have assumed an improper prior for $\theta$, the ex-ante probabilities of bailout are not well defined. A reasonable way to get around this problem is by integrating with respect to a uniform distribution with bounded support, as an attempt to approximate the improper distribution for $\theta$. I assume in my calculations for the probability
of bailout that \( \theta \in [0, 1.12] \). Another interpretation for this exercise is that, although the true distribution for \( \theta \) is a uniform in \([0, 1.12]\), banks are boundedly rational and believe that it is a uniform in \((-\infty, +\infty)\).

In order to interpret figure 9 correctly, we have to keep in mind that the probability of bailout can be modified by changing the approximate distribution of \( \theta \), so the only robust results about bailout probabilities are qualitative and not\(^{22}\) quantitative.

Figure 8 shows the probability of bailout for different values of \( \theta \). Note that with a continuum of small banks, for any value of \( \theta \) less than \( \theta^* \) there is a bailout with certainty. The law of large numbers generates this sharp\(^{23}\) result. When there are large banks, there is distributional uncertainty, making bailouts less likely for low values of \( \theta \) and more likely otherwise. Intuitively, when \( \theta \) is low there is the possibility that large banks receive a signal indicating a high \( \theta \), which would make a bailout less likely\(^{24}\), and vice versa when \( \theta \) is high. With only small banks, the the number of banks with high and low signals about \( \theta \) is balanced by the law of large numbers.

\(^{22}\)Increasing the weight \( w_j \) given to large banks when calculating \( \tilde{L} \) is a simple way to boost the effect of large banks in bailout probabilities. Using a well-defined prior instead of an improper one also helps to generate larger magnitudes.

\(^{23}\)To some extent, this result occurs because the continuum of small banks as a whole has much more information than any large bank.

\(^{24}\)If they receive a lower signal about \( \theta \), the likelihood of bailout would not change, since they were already choosing leverage of 1.
How does the ex-ante probability of bailout change when we make the size of large banks larger, given a fixed number of large banks \( N \)? Integrating figure 8 over \( \theta \) would give us that answer, which is plotted in figure 9. As expected, economies with larger banks tend to have larger bailout probabilities in case of a refinancing shock. Larger banks take more leverage than small banks for most of their private signals, which combined with the inherent uncertainty yield the results in figure 9. Larger values of \( N \) would have shifted the probability of bailout downward in the left plot.

Figure 8: Specification: \( p = 1, \lambda = 0.75, N = 2, \sigma_j = \sigma_k = 1 \)

Figure 9: Specification: \( p = 1, \sigma_j = \sigma_k = 1 \) with \( N = 2 \) in left plot and \( \lambda = 0.75 \) in right plot
Finally, figure 9 also allows us to see how bailout probabilities vary with the number of large banks \(N\), while fixing the overall measure of large banks \(\lambda\). As expected, having smaller large banks decreases the probability of bailout. The mechanisms are the same ones highlighted throughout this section. Assuming a smaller value for \(\lambda\) would have also shifted the probability of bailout downward in the right plot. Figure 9 neatly shows two distinctive empirical predictions of the model.

### 4.3 Asymmetrically informed banks \(\sigma_j < \sigma_k\)

There are different arguments to justify the fact that large banks may have better information than small banks. To capture this hypothesis, I briefly analyze here the case with asymmetric precisions, by assuming \(\sigma_k = 1\) and \(\sigma_k = 2\). This calibration of the model assumes that large banks are relatively better informed about \(\theta\), the fundamentals/type of the central bank.

Figure 10 presents equivalent results to figures 5 and 7.

![Figure 10: Specification: \(p = 1, \lambda = 0.75, N = 2, \sigma_j = 1, \sigma_k = 2\)](image-url)

The optimal leverage function for small banks is flatter than in the case with equally precise signals because, for a given value of \(x_k\), they are more uncertain about the actual value of \(\theta\). However, small banks once again are more aggressive for any value of their private signals when there are large banks present\(^{25}\). The intuition is the same as in the equal precision case. The comparative statics for \(\lambda\) and \(N\) on the probability of bailout are identical to those shown in figure 9 for the case with equal precision.

\(^{25}\)For reasons of space, I present additional plots in the online appendix. The analogous plot to figure 6 makes this last statement clear.
The expected leverage with large and small banks now looks like a combination between the sharper values with a single large bank and the case with only a continuum, which is substantially flatter. Note that expected leverage is larger with a continuum of banks than with large and small banks when \( \theta \) is large. Since for those high values of \( \theta \), bailouts never happen in equilibrium, it is still the case that the probability of bailouts is always larger with large banks. In general we observe a minor decrease in expected leverage for low values of \( \theta \) that is compensated with an increase for high values of \( \theta \). This is driven by the smoother policy followed by both large and small banks.

The online appendix shows equivalent figures for the alternative case, with \( \sigma_j = 2 \) and \( \sigma_k = 1 \). That calibration dampens the optimal leverage choice by large banks, reducing most of the effects highlighted in this paper.

### 4.3.1 Summary of the numerical results

After studying the simulations, we are ready to describe how bank size determines equilibrium outcomes. Robust conclusions are as follows:

- For a given belief about central bank behavior, large banks tend to take more leverage than small banks, although they may actually be more prudent when they expect a tough policy response.

- Small banks take more leverage when large banks are present. This effect is more pronounced when large banks are larger, i.e., when the ratio \( \frac{\lambda}{N} \) is larger. In general, we expect small banks to mimic large bank behavior; because large banks are overall more aggressive than small banks, the actual effect is to make small banks more aggressive. However, if large banks happened to be prudent and small banks were aware of that, they would be more prudent too.

- The expected amount of leverage in the economy is larger whenever large banks are larger.

- The bailout probability is increasing in the measure of large banks, \( \lambda \), given a fixed number \( N \) of large banks.

- The bailout probability is decreasing in the number of large banks, \( N \), given a fixed value for \( \lambda \).

Noisier private signals smooth out all the results.
5 Discussion of the results

5.1 Alternative interpretations and extensions

I have assumed throughout that bank management, acting on behalf of the shareholders, plays a leading role in shaping a bank’s capital structure according to its expectations of government policy. Although plausible, this assumption may seem too crude for some readers. We could instead take a more behavioral approach, and assume that management tends to behave in a myopic way, simply maximizing bank size and consequently its leverage, while discipline is imposed by participants in bank funding markets, in this case, the debtholders. For instance, Shleifer and Vishny (2010) analyze the interaction between securitization and banking through this lens. Figure 11 diagrammatically represents this interpretation.

![Figure 11: An alternative interpretation](image)

The model in this paper can then be reinterpreted such that debtholders are the ultimate decision makers, by modifying the real world counterpart of some parameters and adding a few assumptions. In this interpretation of the world, taking debtholders as a coalition and setting $r_i^K = 0$, $r_i^L$ denotes the net return on loans when there is no shock, $\delta_i$ and $\gamma_i$ are the proportions of the return kept in the crisis state with bailout and without and the quadratic cost term is the debtholders’ cost of distress in crisis states. The main conceptual concern with this interpretation is how to rationalize the choice of the information structure. As is usual in models with endogenous financial markets, information revelation through prices can recover multiplicity of equilibria, as shown by Angeletos and Werning (2006) and Hellwig, Mukherji and Tsyvinski (2006). If there were a single unified financial market that perfectly aggregates information, assuming private signals is not appropriate. However, as long as there exists some kind of market segmentation or noise trading, private signals can be thought of as a modeling device for belief heterogeneity about government policy across different groups of debtholders for different institutions. That would make the argument described by figure 11 feasible.

The model restricts bank decisions to be simultaneous, but in reality banks may gather
some information about the decisions of the others. As long as the information they have is noisy and does not perfectly reveal the fundamentals, we may expect the mechanism described in this paper to hold. This paper does not discuss the possibility of sequential moves, eliminating herding and signaling considerations. In any case, the possibility of herding and signaling should exacerbate the effects of the large banks on outcomes, as discussed in a similar strategic context by Corsetti et al. (2004).

A natural extension of the model can generate endogenous financial crises in environments with a unique bailout authority and multiple independent banking environments. Suppose there are a finite number of countries, each with an independent banking sector but a single bailout authority. In that case, if systemic shocks come in a sequential order, banks are able to learn about the type of the bailout authority $\theta$. Under the interpretation of the model that makes financial markets the actual decision makers, every time that the central bank makes a bailout decision, financial markets update their beliefs over $\theta$, shifting optimal leverage choices and shaking debt markets.

This extension can accurately describe the European Union response to the Greek bailout and the concern created in financial markets about other countries like Ireland, Portugal or Spain. A serious analysis of that environment is not the purpose of this paper, but clearly the central bank should have a more meaningful objective function that takes into consideration learning and dynamics in that case.

Lastly, even though this paper is framed in terms of leverage choices, a similar analytical framework could be used to describe lobbying activities. The importance of lobbying has been stressed in recent work by Acemoglu et al. (2010) and Johnson and Kwak (2010), who give empirical and anecdotal support to the view that large financial institutions have been able to internalize government behavior by assuming more risks. An appropriate reinterpretation of $L_i$ in that context would be as the amount of resources spent on lobbying by each institution. Leaving lobbying aside, that branch of work provides evidence to suggest that bank management may behave with the same degree of sophistication assumed in the main exposition of this paper.

5.2 Empirical relevance

The increasing concentration of the U.S. banking industry in the last fifty years, as documented by Janicki and Prescott (2006), is a well established empirical fact; the authors also emphasize that the right tail of the bank size distribution features too many large banks to fit lognormal or Pareto distributions. Other countries show similar patterns that support the
idea that deregulation has increased concentration. The theory proposed in this paper can be understood as a way to rationalize the endogenous choice of size by banks. If, controlling for other factors, banks are aware that their size can help them induce a more favorable policy response in crisis states, they will prefer to be larger or to merge ex-ante. This mechanism may not be the only cause of the trend in banking concentration, but it is likely to be a contributing factor, especially for the very large banks that are clearly aware of their size implications.

This model presents clear testable predictions. Controlling for other factors, figures (7) and (9) imply that economies with larger banks will take more risk and that increasing the size of large entities or allowing for a few powerful banks will make bailouts more likely. Unfortunately, as discussed in the survey by Gorton and Winton (2003), testing moral hazard hypotheses is fraught with difficulties. An ideal test of this model, although theoretically feasible, may be extremely complicated to carry out in practice. A clear experimental design would generate an exogenous change in the size of the banking industry, while keeping constant the rest of the environment, primarily the type $\theta$ of the central bank. The model would also predict that, when banking is more concentrated, large banks take higher leverage and small banks do the same in a more moderate way.

Looking at market variables to test the implications of the model would introduce further complications. An analysis of CDS spreads would imply that, controlling for other factors, we should expect smaller CDS spreads in banking systems with a large number of larger banks, since the model predicts a larger likelihood of bailout. For many countries, “too big to save” problems, that are not discussed in this paper, may also arise when the size of the banking sector exceeds a certain fraction of national GDP. Overall, I believe that this paper provides a simple but realistic framework to further analyze these issues.

5.3 Policy implications

The welfare implications of this model are not straightforward. In the model, there is a strategic complementary in leverage decisions due to the policy response but, as studied in detail by Angeletos and Pavan (2007), strategic complementarities and welfare need not be linked. The fact that I have abstracted from providing a truly detailed microfoundation does not help with this dimension.

Consequently, to address policy issues, we need to take a stand on whether banks acting in a decentralized environment fully internalize their social costs or not. If bailouts are costly, either because they involve raising distortionary taxes or because they distort intertemporal substitution through low interest rates, as in Farhi and Tirole (2010), there is a clear rationale
for curtailing leverage. Moreover, fire sales externalities or increased adverse selection may suggest that leverage is socially suboptimal even without bailouts. According to either view, since large banks increase both the expected amount of leverage and the probability of bailout, they should be subject to special screening.

If the government could commit\(^{26}\) to a value of \(\theta\) to reduce leverage, it would choose the highest possible, committing not to bail out any large or small bank. Accordingly, it would also try to minimize \(\delta_i - \gamma_i\), the difference in recovery rates in the crisis state when a bank is bailed out or not. The model assumes that because of time inconsistency problems, such as to those discussed by Chari and Kehoe (2010), these ex-post policies are not feasible.

Nonetheless, there are some ex-ante policies that may be effective. Imposing capital requirements or even setting a direct cap on size can implement any desired output in the model. Using a direct cap on size is a policy that has recently received special attention and it arises as a natural policy in this model: forcing all banks to be small would obviously kill any additional leverage due to size internalization. Increased capital requirements only for large banks is another policy that emerges from this model; this would curtail the large banks’ leverage and, in equilibrium, would decrease the spillover effects for small banks.

Finally, although this paper clearly points out a theoretical argument against size concentration in the banking sector, other defensible arguments exist in favor of large financial institutions. Better diversification and risk management, economies of scale, ability to handle large projects, creation of insensitive claims à la Gorton and Pennacchi (1990) or reduction of international frictions as in Freixas and Holthausen (2005) are relevant arguments on the opposite side of the debate. Trading off these considerations will determine the optimal size distribution for the banking sector.

6 Conclusion

This paper shows that the size distribution of financial institutions is crucial for the ex-ante determination of leverage when bailouts are possible. Large banks are able to internalize the effect of their size on the likelihood of bailout and thus increase their leverage in equilibrium. Their increased exposure creates spillovers for small banks, which also raise their leverage, since they are shielded by the behavior of large banks. Consequently, the expected amount of leverage in the banking sector and the probability of bailouts are larger with large banks than

\(^{26}\)The full analysis when the government can affect \(\theta\) is far from trivial. See Angeletos, Hellwig and Pavan (2006) for the issues that may arise then.
with only small banks.

The mechanism described in this paper relies on minimal assumptions about complementarities created by the policy response and is robust to different theories of capital structure. No preferential treatment by the policymaker of large versus small entities and herding or signaling considerations are part of this model. Those effects would simply amplify my conclusions.

My results show that the “too many too fail” critique to the “too big too fail” problem is not well grounded. The model gives support to the idea that regulators and policymakers must pay special attention to large financial institutions, since they have a direct motive to take more risk and also because their behavior influences the decisions of small banks in equilibrium. In this model, a regulator that closely monitors and restricts funding decisions of large financial institutions arises as a natural policy.
APPENDIX A: Sufficient conditions for uniqueness

This appendix shows that high values of $\sigma_j$ and $\sigma_k$ are sufficient conditions for equilibrium uniqueness. For simplicity, I analyze the case discussed in the numerical simulations, i.e., $\gamma_i = 0$, $\delta_i = 1$, $r_i^K = 0$ and $r_i^L = p\kappa_i$; the reasoning for the general case is identical. The equilibrium optimal leverage choices can be seen as a system of functional equations. Showing that the optimal leverage choices define a contraction in a suitable space is enough to show that the equilibrium is unique, by direct application of the Contraction Mapping Theorem. I first define an operator $T$ in that suitable space and then provide sufficient conditions to guarantee that $T$ is a contraction. Since the operator is vector valued, I need to use a slightly more general version than the usual Blackwell’s conditions as stated, for instance, in Stokey, Lucas and Prescott (1989), which are defined only for maps from $\mathbb{R}^m \to \mathbb{R}$.

I denote by $S$ the space of bounded continuous functions $f : \mathbb{R}^2 \to \mathbb{R}^2$ with the metric $\sup_{x \in \mathbb{R}^2} \{\max \{|f_1 + g_1|, |f_2 + g_2|\}\}$. For a function $f$ in that space, I write $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$.

I define $T : S \to S$ as a functional operator from the space $S$ into itself: $(TL)(x) \equiv \begin{pmatrix} (TL_k)(x_k) \\ (TL_j)(x_j) \end{pmatrix}$ with

$$(TL_k)(x_k) \equiv 1 - p + p F_{S_{L_k,L_j}|x_k}(0)$$
$$(TL_j)(x_j) \equiv 1 - p + p \left( F_{S_{L_k,L_j}|x_j} \left( \frac{\lambda}{N} L_j(x_j) \right) + L_j(x_j) \frac{\lambda}{N} f_{S_{L_k,L_j}|x_j} \left( \frac{\lambda}{N} L_j(x_j) \right) \right)$$

And

$$F_{S_{L_k,L_j}|k}(0) = \mathbb{E}_k \left[ 1 \left( \theta - (1 - \lambda) L(\theta) - \sum_{j=1}^N \frac{\lambda}{N} L_j(x_j) \leq 0 \right) \right]$$

$$F_{S_{L_k,L_j}|x_j} \left( \frac{\lambda}{N} L_j(x_j) \right) = \mathbb{E}_j \left[ 1 \left( \theta - (1 - \lambda) L(\theta) - \sum_{j' \neq j} \frac{\lambda}{N} L_{j'}(x_{j'}) \leq \frac{\lambda}{N} L_j(x_j) \right) \right]$$

$L(\theta)$ is defined as $\int L_k dk \to_p (1 - \lambda) L(\theta)$. The pdf corresponding to $F_{S_{L_k,L_j}|x_j}$ is denoted by $f_{S_{L_k,L_j}|x_j}$; this density is well defined since the conditional distribution of $\theta|x_i$ and $x_i'|x_i$ is normal, $L_k$ is strictly decreasing and consequently invertible and, when inverting $L_j$, there is only a finite number of roots.$^{27}$ Given assumption 1 in the text, both large and small banks solve a strictly concave for each $x_i$, so an application of the Theorem of the Maximum guarantees continuity of $L_k$ and $L_j$. Since the limit of $L_i$ with a signal $x_i \to \infty$ is 1 and with $x_i \to -\infty$ is 0, boundedness is also guaranteed.

$^{27}$See Casella and Berger (2001).
Fact 1. **Generalized Blackwell’s sufficient conditions:** Let $X \subseteq \mathbb{R}^m$ and $B^2(X)$ be the space of bounded functions $f : X \rightarrow \mathbb{R}^2$ with the the norm defined by $\|f\| = \sup_{x \in \mathbb{R}^m} \{\max \{|f_1|, |f_2|\}\}$. Let $T : B^2(X) \rightarrow B^2(X)$ be an operator satisfying:

1. **Monotonicity:** If $g,h \in B^2(X)$ are such that $g(x) \leq h(x)$ for all $x \in X$, then $(Tg)(x) \leq (Th)(x)$ for all $x \in X$.

2. **Discounting:** Let the function $g + a$, for $g \in B^2(X)$ and $a \in \mathbb{R}_+$ be defined by $(g + a)(x) = g(x) + a$. There exists $\beta \in (0,1)$ such that for all $g$ and $a$ and all $x \in X$, $[T(g + a)](x) \leq [Tg](x) + \beta a$.

If both conditions hold, the operator $T$ is a contraction.

*Proof:* See online appendix. The proof is a straightforward generalization of the analogous one in Stokey, Lucas and Prescott (1989).

**Claim 1.** If the private signals are noisy enough, this is, $\sigma_k$ and $\sigma_j$ are large enough, there is a unique equilibrium.

*Proof:* Given claim 1, it is sufficient to show monotonicity and discounting for $T$.

**Monotonicity:** Consider two functions $g \leq h$. Monotonicity holds if $(Tg) - (Th) \leq 0$.

For the small bank case we need that

$$(Tg_1) - (Th_1) = p \left[ F_{S_g|x_j}(0) - F_{S_h|x_j}(0) \right] \leq 0$$

For this condition hold, it is sufficient to show that $F_{S_g|x_j}$ first-order stochastically dominates $F_{S_h|x_j}$, which trivially holds given the definition of $S$. Intuitively, the existence of strategic complementarities makes the operator $T$ monotonic.

For the large bank case we need the equivalent condition

$$(Tg_2) - (Th_2) = p \left[ \left( F_{S_g|x_j} \left( \frac{\lambda}{N} g_2 \right) - F_{S_h|x_j} \left( \frac{\lambda}{N} h_2 \right) \right) + \frac{\lambda}{N} \left( g_2 f_{S_g|x_j} \left( \frac{\lambda}{N} g_2 \right) - h_2 f_{S_h|x_j} \left( \frac{\lambda}{N} h_2 \right) \right) \right] \leq 0$$

Using an analogous argument to the case with only small banks, it is easy to see that $F_{S_g|x_j}$ stochastically dominates $F_{S_h|x_j}$, what implies that $F_{S_g|x_j} \left( \frac{\lambda}{N} g \right) \leq F_{S_h|x_j} \left( \frac{\lambda}{N} g \right) \leq F_{S_h|x_j} \left( \frac{\lambda}{N} h \right)$, where the last inequality follows from the monotonicity of a cdf. This term is again driven by strategic complementarities. The sign of the second term $g_2 f_{S_g|x_j} \left( \frac{\lambda}{N} g \right) - h_2 f_{S_h|x_j} \left( \frac{\lambda}{N} h \right)$ cannot be unambiguously determined, but when the private signals have enough noise, $f_{S_g|x_j} \approx f_{S_h|x_j} \approx 0$ and this last term is small compared to the one determined by the difference
between cdf’s. Intuitively, it may be the case that a large bank wants to be less aggressive when small banks or other large banks are taking more leverage; if there is enough noise, this strategic substitutability is never strong enough. I have checked the results numerically for the calibrations used in the paper and \((Tg_2) - (Th_2)\) is monotonic in the relevant region for different choices of \(g\) and \(h\). For reference, if \(F\) happened to be a normal \(N(0, \sigma)\), \(xf(x)\) would be increasing in the interval \([-\sigma, \sigma]\); plotting both the cdf and pdf for different values of \(\sigma\) is helpful to understand this result.

**Discounting:** We must show that there exists \(\beta \in (0, 1)\) such that for all \(g, a \geq 0\) and all \(x \in X\), \(\frac{T(g+a) - T(g)}{a} \leq \beta\) holds. Let’s work first with \(g_1\)

\[
\frac{T(g_1 + a) - T(g_1)}{a} = p \left( \frac{F_{Sg+a|x_k}(0) - F_{Sg|x_k}(0)}{a} \right) = p \left( \frac{F_{Sg|x_k}(a) - F_{Sg|x_k}(0)}{a} \right) < 1
\]

\[
\lim_{a \to 0} \frac{T(g_1 + a) - T(g_1)}{a} = p \left( f_{Sg|x_k} \left( \frac{\lambda}{N} g_2 \right) \right) < 1
\]

By taking the limit with \(a \to 0\), it is a sufficient condition that \(f_{S|x_j} < 1\). Enough noise in the private signals \(\sigma_k\) is a sufficient condition for this to hold. Let’s proceed with \(g_2\)

\[
\frac{T(g_2 + a) - T(g_2)}{a} = p \left( F_{Sg+a|x_j} \left( \frac{\lambda}{N} (g_2 + a) \right) - F_{Sg|x_j} \left( \frac{\lambda}{N} g_2 \right) \right) + \frac{\lambda}{N} g_2 \left( f_{Sg+a|x_j} \left( \frac{\lambda}{N} (g_2 + a) \right) - f_{Sg|x_j} \left( \frac{\lambda}{N} g_2 \right) \right)
\]

\[
+ p \frac{\lambda}{N} f_{Sg+a|x_j} \left( \frac{\lambda}{N} (g_2 + a) \right)< 1
\]

By taking the limit with \(a \to 0\), we can write

\[
\lim_{a \to 0} \frac{T(g_2 + a) - T(g_2)}{a} = p \left( f_{Sg|x_j} \left( \frac{\lambda}{N} g_2 \right) \right) + \frac{\lambda}{N} g_2 f_{Sg|x_j} \left( \frac{\lambda}{N} g_2 \right) + \frac{\lambda}{N} f_{Sg|x_j} \left( \frac{\lambda}{N} g_2 \right) < 1
\]

\(^{28}\)I write “happened to be” because, in this model, \(S\) need not be normally distributed. When solving the model numerically, it actually fails the Kolmogorov-Smirnov test. However, a plot of the cdf of \(S\) is not distinguishable from the normal cdf by mere observation.
In this case we need that both \( f_{S|\sigma_j} \) and \( f'_{S|\sigma_j} \) small enough. Again, enough noise in the private signals \( \sigma_j \) is a sufficient condition for these to hold.

The conditions discussed in this appendix are only sufficient but not necessary to have a unique equilibrium. For instance, in the case with a continuum a small banks, the model fits with minor modifications\(^{29}\) into the general framework of Frankel, Morris and Pauzner (2003). Any value of \( \sigma_k \) delivers a unique equilibrium in that case. Since showing uniqueness for the general case would require major modifications to their proof\(^{30}\), I have opted for a more direct approach that only imposes the assumption of enough noise. Notice also that enough noise is necessary to have a convex problem for large banks. Since high levels of uncertainty about policy responses are a distinctive feature of this economic environment, these conditions have a strong empirical support.

\(^{29}\) Basically, the signals should be modified to have bounded support. We could think in that case that the normal provides an approximation to a distribution with bounded support.

\(^{30}\) My model with large agents and vanishing noise does not satisfy strategic complementarities for all the parameter space, state monotonicity, payoff continuity and bounded derivatives.
References


Complete Motivational Quotes

Breaking up big banks wouldn’t really solve our problems, because it’s perfectly possible to have a financial crisis that mainly takes the form of a run on smaller institutions. (...) Breaking up big financial institutions wouldn’t prevent future crises, nor would it eliminate the need for bailouts when those crises happen. The next bailout wouldn’t be concentrated on a few big companies - but it would be a bailout all the same. I don’t have any love for financial giants, but I just don’t believe that breaking them up solves the key problem.


But, while regulation must address the oversized bank balance sheets that were at the root of the crisis, the IMF is right not to focus excessively on fixing the “too big to fail” problem. A surprising number of pundits seem to think that if one could only break up the big banks, governments would be far more resilient to bailouts, and the whole “moral hazard” problem would be muted. That logic is dubious, given how many similar crises have hit widely differing systems over the centuries. A systemic crisis that simultaneously hits a large number of medium-sized banks would put just as much pressure on governments to bail out the system as would a crisis that hits a couple of large banks.

Kenneth Rogoff. All for One Tax and One Tax for All? Project Syndicate, 2010-04-29

Most observers who study this believe that to try to break banks up into a lot of little pieces would hurt our ability to serve large companies and hurt the competitiveness of the United States.

But that’s not the important issue. They believe that it would actually make us less stable, because the individual banks would be less diversified and, therefore, at greater risk of failing, because they would haven’t profits in one area to turn to when a different area got in trouble.

And most observers believe that dealing with the simultaneous failure of many small institutions would actually generate more need for bailouts and reliance on taxpayers than the current economic environment.

Lawrence Summers. Interview with Jeffrey Brown, PBS Newshour, 2010-04-22
1 Numerical procedures

I use a standard collocation method\(^1\) to find the optimal leverage functions in equilibrium. After experimenting with both Chebyshev polynomials and splines, quadratic or cubic splines seem to be the basis functions that provide a fast and accurate approximation. Quadratic splines provide roughly the same accuracy as cubic splines but reduce substantially the execution time, so I use them for all the figures of the general case presented in the paper.

When solving for the equilibrium fixed point, I approximate \(L_k, L_j\) and \(L(\theta)\). Approximating this last function is in theory not necessary, but making it an equilibrium object instead of having to calculate it from the approximation of \(L_k\) every time considerably reduces the execution time. I use between 50 and 70 Chebyshev nodes for both \(L_k\) and \(L_k\) and 20 Chebyshev nodes for \(L(\theta)\). Changing the number of nodes doesn’t change the results. For the single large bank case, I use 300 Chebyshev nodes.

When possible, I use Gauss-Hermite quadrature to take expectations of normally distributed random variables. Unfortunately, the distribution \(S|x_j\) is not normal, so I must approximate it numerically using Monte Carlo methods, which drastically increases the execution time. Lastly, I rely on Broyden’s method for solving the nonlinear system of equations in the collocation problem. The code can be easily modified to use Matlab’s \texttt{fsolve}\(^2\).

In order to replicate the figures in the paper, follow the instructions in the documentation file. That file also discusses the convenience of using different parameters. In order to quickly generate the figures where \(N\) and \(\lambda\) vary, I use Matlab Parallel Toolbox capabilities. The running time for the general case is between 3500 to 7000 seconds in a 2.3GHz CPU with 4 GB of RAM, depending on the parameter configuration.

2 Forming conditional beliefs

When a bank observes a signal \(x_i\), it must form beliefs about the actual value of \(\theta\) and the signals of the other banks in order to compute the distribution for \(F_{S|x_j}\) and \(f_{S|x_j}\).

Given the information structure assumed in the paper, the joint conditional distribution of beliefs given a signal \(x_i\) is

\[
\begin{pmatrix}
\theta \\
x_j \\
x_k
\end{pmatrix}
\sim N\left(
\begin{pmatrix}
x_i \\
x_i \\
x_i
\end{pmatrix},
\begin{pmatrix}
\sigma_i^2 & \sigma_i^2 & \sigma_i^2 \\
\sigma_i^2 & \sigma_i^2 + \sigma_i^2 & \sigma_i^2 \\
\sigma_i^2 & \sigma_i^2 & \sigma_i^2 + \sigma_i^2 + \sigma_i^2
\end{pmatrix}\right)
\]

The variance of \(\theta\) only depends on the noise of the own signal, as well as the covariances. The variances depend on both the own noise and the noise of the other banks.

When there is a prior/public signal \(y\), as described in section 3 in this appendix, the conditional distribution for \(\theta\) given \(y\) and \(x_i\) is

\[
\theta|y,x_i \sim N\left(\frac{1}{\sigma_y^2 + \frac{1}{\sigma_i^2}} y + \frac{1}{\sigma_y^2 + \frac{1}{\sigma_i^2}} x_i, \left(\frac{1}{\sigma_y^2} + \frac{1}{\sigma_i^2}\right)^{-1}\right)
\]

The joint conditional distribution then becomes

\(^1\)Judd (1998) and Miranda and Fackler (2004) are standard references.
3 Extension with vulture behavior

As discussed in the paper, allowing government policies to modify the value of $\gamma_i$ and $\delta_i$, can strongly affect the actual equilibrium outcomes. More specifically, in those situations in which a systemic shocks occurs and the government does not step in, the possibility of strong banks taking over distressed banks gives a rationale to reduce leverage ex-ante. The papers by Perotti and Suarez (2002) and Acharya and Yorulmazer (2007) analyze similar situations. This fact introduces a source of strategic substitutability that leans against the main mechanism described in the model. Here, I present a simple extension to understand that mechanism inside the model.

More generally, we could make government policy contingent on individual behavior. Any extension along these lines would modify optimal responses without changing the mechanism of the model.

The only difference with respect to the main version of the paper is that banks now take into consideration that $\gamma_i$, the proportional loss rate in the bad state, now depends on the leverage they assume. To have a better intuition, I assume here that $\gamma_i (L_i) = \gamma_i^0 - \tau L_i$, with $\gamma_i^0 > \tau \frac{\delta_k}{p K_i}$. In other words, in the no bailout state, the recovery rate for each bank depends negatively in the amount of leverage taken. The case with $\tau = 0$ is the one described in the main text.

The problem for each bank now becomes

$$\max_{L_i} (1 - p) r_i^L L_i + p \left[ \mathbb{E}_i \left[ q (\theta, \bar{L}) \right] \delta_i r_i^L L_i + (1 - \mathbb{E}_i \left[ q (\theta, \bar{L}) \right] ) \gamma_i (L_i) r_i^L L_i - \frac{\kappa_i L_i^2}{2} \right] - r_i^K L_i$$

and the first order conditions that characterize the equilibrium\footnote{As in the general case, enough noise and a smooth $\gamma'$ are needed to ensure convexity of the individual problems.} are

$$L_i^* = \left\{ 1 - p + p \left( \gamma_i (L_i) + [\delta_i - \gamma_i (L_i)] + \mathbb{E}_i \left[ q (\theta, \bar{L}) \right] + \frac{\partial \mathbb{E}_i [q (\theta, \bar{L})]}{\partial L_i} L_i \right) + (1 - \mathbb{E}_i \left[ q (\theta, \bar{L}) \right] ) \gamma_i' (L_i) L_i \right\} \frac{r_i^L - r_i^K}{p \kappa_i}$$

Substituting $\gamma_i (L_i) = \gamma_i^0 - \tau L_i$ and assuming $r_i^K = p \kappa_i$, $p = 1$ and $r_i^K = 0$

$$L_i^* = \frac{1}{1 + \tau (1 - \mathbb{E}_i \left[ q (\theta, \bar{L}) \right] )} \left\{ \gamma_i^0 - \tau L_i + [\delta_i - (\gamma_i^0 - \tau L_i)] + \mathbb{E}_i \left[ q (\theta, \bar{L}) \right] + \frac{\partial \mathbb{E}_i [q (\theta, \bar{L})]}{\partial L_i} L_i \right\}$$

To get some intuition, let’s take limits in $\mathbb{E}_i \left[ q (\theta, \bar{L}) \right]$, and assume that $\frac{\partial \mathbb{E}_i [q (\theta, \bar{L})]}{\partial L_i} = 0$
\[
\lim_{E_i[q(\theta, L)] \to 0} L_i^* = \frac{1}{1 + 2\tau_i \gamma_i} \\
\lim_{E_i[q(\theta, L)] \to 1} L_i^* = 1
\]

These two limit cases provide sufficient intuition for the general case. Since \(\tau > 0\), banks are more conservative than in the baseline case with \(\tau = 0\) when they expect there not to be a bailout. I have made the assumption, that only \(\gamma_i\) is increasing in the conservativeness of leverage decisions. If we think that bailouts are supposed to be targeted only at institutions with healthier balance sheets, an analogous effect for bailout states can be create by endogenizing \(\delta_i\).

4 Extension with an informative prior/public signal

Here I analyze what happens when all banks share an informative prior/public signal. The information structure in the main version of the paper consists of a type \(\theta\) for the central bank drawn from an improper prior in the real line and private signals \(x_k, x_j\), given by

\[
x_k = \theta + \sigma_k \varepsilon_x \\
x_j = \theta + \sigma_j \varepsilon_x
\]

Instead of assuming that \(\theta \sim U[-\infty, +\infty]\), we can be more general and assume that all banks share a prior/public signal \(y\) about the fundamentals given by

\[
y = \theta + \sigma_y \varepsilon_x
\]

where all the noise terms \(\varepsilon_x\) are distributed iid standard normal. The leading case analyzed in the paper corresponds to the limiting case when \(\sigma_y \to \infty\).

As usual in this literature, I define \(\alpha_k = \frac{1}{\sigma_k^2}\), \(\alpha_j = \frac{1}{\sigma_j^2}\), \(\alpha_y = \frac{1}{\sigma_y^2}\) as the precisions of each signal. I alternate between notations when convenient. With this public signal, the signal extraction problem that each bank faces implies that\(^3\)

\[
\theta|y, x_i \sim N\left(\frac{\alpha_y y + \alpha_i x_i}{\alpha_y + \alpha_i}, (\alpha_y + \alpha_i)^{-1}\right)
\]

See section 3 in this appendix for the joint distribution of \(\theta, x_k\) and \(x_j\) given \(x_i\).

Here I only discuss the case with a continuum of small banks \(\lambda = 0\). The argument used to derive sufficient conditions for uniqueness in the general case can be applied identically to this situation, with the exception that now \(L_j\) and \(L_k\) are conditioned on \(y\), and all the updating must take this into account. As in the rest of the paper, a non-precise public signal is necessary because it prevents the coordination of small banks and dampens nonconvexities in the large banks' problem.

\(^3\)See, for instance, the appendix in Vives (2008).
4.1 No private information with only small banks

Without private information and with common knowledge of the public signal, the solution now depends on the signal $y$. If $\sigma_y$ is too precise, we go back to the common knowledge intuition where we have multiplicity. If $\sigma_j$ is too imprecise, we can recover uniqueness. This is intuitive, because a noisy public signal reduces complementarities. This problem is identical to the individual problem of the myopic large bank that doesn’t internalize its size: in this case, all small banks share the same information and therefore play identically, but they cannot fully coordinate their actions.

I define $\tilde{L}(y) = \int L_k(y) \, dk$. Hence, in equilibrium

$$\tilde{L}(y) = \Phi \left[ \frac{\tilde{L}(y) - y}{\sigma_y} \right]$$  \hspace{1cm} (1)

For any $y$, we know that $\tilde{L}(y)$ is continuous, strictly increasing in $\tilde{L}(y)$, $\lim_{\tilde{L}(y) \to \infty} \Phi \left[ \frac{\tilde{L}(y) - y}{\sigma_y} \right] = 1$ and $\lim_{\tilde{L}(y) \to -\infty} \Phi \left[ \frac{\tilde{L}(y) - y}{\sigma_y} \right] = 0$, which guarantees existence. For uniqueness, we have to check that the slope of the RHS is less than one.

Taking the derivative of the RHS in (1), we find the condition for uniqueness

$$\max \frac{1}{\sigma_y} \Phi \left[ \frac{\tilde{L}(\theta) - \theta}{\sigma_y} \right] = \frac{1}{\sigma_y} \frac{1}{\sqrt{2\pi}} < 1 \Rightarrow \frac{1}{\sqrt{2\pi}} < \sigma_y$$  \hspace{1cm} (2)

Therefore, if $\sigma_y$ is too small (the precision of the public information is too large), there are multiple equilibria. This is not surprising, since the common knowledge case is $\sigma_y \to 0$. For comparison below, (2) can be written as $\alpha_y < 2\pi$.

4.2 Public and private information

The optimal policy for each agent in the continuum when $\lambda = 0$ is given by

$$L_k^*(x_k, y) = 1 - p + p \Phi \left( \sqrt{\alpha_y + \alpha_k} \left( \tilde{L}(\theta, y) - \frac{\alpha_y}{\alpha_y + \alpha_k} y - \frac{\alpha_k}{\alpha_y + \alpha_k} x_k \right) \right)$$

In order for an equilibrium to be characterized in this case, we need to find a $\theta^*$ such that

$$\theta^* = \tilde{L}(\theta^*, y) = 1 - p + p \int \Phi \left( \sqrt{\alpha_y + \alpha_k} \left( \theta^* - \frac{\alpha_y}{\alpha_y + \alpha_k} y - \frac{\alpha_k}{\alpha_y + \alpha_k} x_k \right) \right) d\Phi \left( \frac{x_k - \theta^*}{\sigma_k} \right)$$

Or

$$\theta^* = 1 - p + p \int \Phi \left( \sqrt{\alpha_y + \alpha_k} \left( \theta^* - \frac{\alpha_y}{\alpha_y + \alpha_k} y - \frac{\alpha_k}{\alpha_y + \alpha_k} x_k \right) \right) \frac{1}{\sigma_k} \phi \left( \frac{x_k - \theta^*}{\sigma_k} \right) \, dx_k$$  \hspace{1cm} (3)

It can be shown (see section 5 in this appendix) that the right-hand side in (3) can be written as

$$F(\theta^*) = \Phi \left( \frac{\alpha_y (\theta^* - y)}{\sqrt{\alpha_y + 2\alpha_k}} \right)$$

In order to have a unique equilibrium, $F'(\theta^*)$ must be less than unity. A sufficient condition is
max \( F'(\theta^*) = \max \frac{\alpha_y}{\sqrt{\alpha_y + 2\alpha_k}} \phi \left( \frac{\alpha_y (\theta^* - y)}{\sqrt{\alpha_y + 2\alpha_k}} \right) < 1 \) (4)

\[
\frac{\alpha_y}{\sqrt{\alpha_y + 2\alpha_k}} < \sqrt{2\pi}
\]

Rearranging

Since \( \alpha_y \) and \( \alpha_k \) are positive, for a given value of \( \alpha_k \), there is always a region such that, under high precision of the public signal there is multiplicity and for low values of the public signal there is uniqueness. The particular case studied in the paper, \( \alpha_y = 0 \), clearly satisfies condition (6).

As expected, when the precision of public information is large enough with respect to the precision of private information, multiple equilibria are possible. Note that the case with a single large bank and a public signal is identical to the case with a private signal, with the exception that now the single large bank has two different signals and more precise information.

4.3 Smoothing the policy response

Observe that allowing for a public signal is analogous to a formulation in which the central bank makes small mistakes about the actual value of \( \tilde{L} - \theta \). In other words, from a strategic perspective, it doesn’t matter whether \( \theta \) is imperfectly observed or the government policy includes noise ex post.

In the main model we have

\[
q (\theta, \tilde{L}) = \begin{cases} 
1, & \tilde{L} - \theta \geq 0 \\
0, & \tilde{L} - \theta < 0
\end{cases}
\]

but we can also think that the policy can be executed according to

\[
q (\theta, \tilde{L}) = \Phi_{\sigma_y} [\tilde{L} - \theta]
\]

where \( \Phi \) is the cdf of a normal with mean \( \mu = 0 \) and standard deviation \( \sigma_y \geq 0 \). In this case, when \( \sigma_y \to 0 \) we go back to usual formulation in the paper. As we anticipated, a noisy policy response (high \( \sigma_y \)), breaks down the strategic complementarities, eliminating the multiplicity of equilibrium.

Intuitively, an imprecise public signal reduces the strategic complementarities. When the public signal is arbitrarily precise, we are back to the common knowledge case.

5 Conditions for convexity in the large bank problem

The problem for a given large bank \( j \) is:

\[
\max_{L_i} (1 - p) L_i + p \left( \mathbb{E}_i \left[ q (\theta, \tilde{L}) \left[ \delta_i L_i - \frac{L_i^2}{2} \right] + (1 - \mathbb{E}_i \left[ q (\theta, \tilde{L}) \right]) \left[ \gamma_i A_i - \frac{L_i^2}{2} \right] \right) \right)
\]
When there is a single large bank, \( \lambda = 1 \) and \( N = 1 \), its first order condition is

\[
L_j^* (x_j) = 1 - p + p \left( \Phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) + L_j^* (x_j) \frac{1}{\sigma_j} \phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) \right)
\] (5)

In general, equation (5) is necessary but not sufficient to characterize the solution to this problem. For each \( x_j \), a sufficient condition for the single large bank problem to have a unique solution is that the slope of right-hand side in (5) is less that unity, that is, that the problem is strictly concave. We are simply checking the second order conditions.

We need that

\[
\max_{x_j} \left\{ p \left( \frac{1}{\sigma_j} \phi \left[ \frac{L_j^* (x_j) - x_j}{\sigma_j} \right] \right) + \frac{1}{\sigma_j} \phi \left[ \frac{L_j^* (x_j) - x_j}{\sigma_j} \right] - L_j^* (x_j) \frac{1}{\sigma_j^2} \phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) \phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) \right\} < 1
\]

\[
\max_{x_j} \left\{ p \left( \frac{2}{\sigma_j} \phi \left[ \frac{L_j^* (x_j) - x_j}{\sigma_j} \right] \right) - L_j^* (x_j) \frac{1}{\sigma_j^2} \phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) \phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) \right\} < 1
\]

\[
p \frac{1}{\sqrt{2\pi}} \left( \frac{2}{\sigma_j} - L_j^* (x_j) \frac{1}{\sigma_j^2} \phi \left( \frac{L_j^* (x_j) - x_j}{\sigma_j} \right) \right) < 1
\]

(6)

It is clear from (6) that a large value of \( \sigma_j \) is sufficient for this condition to hold.

The equivalent first order condition to (5) in the general case with \( \lambda \in [0, 1] \) is

\[
L_j^* = 1 - p + p \left( \gamma_i + (\delta_i - \gamma_i) \left[ E_i \left[ q (\theta, \tilde{L}_j) \right] + L_j^* \frac{\partial E_i}{\partial L_i} \left[ q (\theta, \tilde{L}_j) \right] \right] \right)
\]

And the analogous condition to (6)

\[
2 \frac{\partial E_i}{\partial L_i} \left[ q (\theta, \tilde{L}_j) \right] + L_j^* \frac{\partial^2 E_i}{\partial L_i^2} \left[ q (\theta, \tilde{L}_j) \right] < 1
\]

It is clear that \( \sigma_j \) large enough is sufficient to have a well defined concave problem. When large banks are uncertain about \( \theta \) and the actions of the other players, both \( \frac{\partial E_i}{\partial L_i} [q(\theta, \tilde{L})] \) and \( \frac{\partial^2 E_i}{\partial L_i^2} [q(\theta, \tilde{L})] \) cannot be too large. When this assumption doesn’t hold, the problem still has a unique global maximum for almost all values of \( x_j \), but then the policy functions are discontinuous and the analysis is much more cumbersome.

6 Auxiliary technical results

**Definition 1.** A function \( T \) that maps \( T : S \to S \), with \( (S, d) \) being a metric space is a **contraction** if there exists \( \beta \in [0, 1] \) such that

\[
d(Tx, Ty) \leq \beta d(x, y) \quad \forall x, y \in S
\]

**Lemma 1.** Generalized Blackwell’s sufficient conditions: Let \( X \subseteq \mathbb{R}^m \) and \( B^2 (X) \) be the space of bounded functions \( f : X \to \mathbb{R}^2 \) with the the norm defined by \( \|f\| = \sup_{x \in \mathbb{R}^m} \{\max \{|f_1|, |f_2|\} \} \). Let \( T : B^2 (X) \to B^2 (X) \) be an operator satisfying:
1. **Monotonicity:** If \( g, h \in B^2(X) \) are such that \( g(x) \leq h(x) \) for all \( x \in X \), then \( (Tg)(x) \leq (Th)(x) \) for all \( x \in X \).

2. **Discounting:** Let the function \( g + a \), for \( g \in B^2(X) \) and \( a \in \mathbb{R}_+ \) be defined by \( (g + a)(x) = g(x) + a \).

There exists \( \beta \in (0,1) \) such that for all \( g \) and \( a \) and all \( x \in X \), \([T(g + a)](x) \leq [Tg](x) + \beta a\).

If both conditions hold, the operator \( T \) is a contraction.

**Proof.** If \( f, g \in B^2(x) \) are such that \( f_1(x) \leq g_1(x) \) and \( f_2(x) \leq g_2(x) \) for all \( x \in X \), then we can write \( f \leq g \). Fix \( x \in X \). Define \( f(x) - g(x) \equiv \left( \begin{array}{c} f_1(x) - g_1(x) \\ f_2(x) - g_2(x) \end{array} \right) < 0 \). Then \( f(x) - g(x) \leq \sup_{x \in \mathbb{R}^m} \{\max \{|f_1 - g_1|, |f_2 - g_2|\}\} \), and this is true for any \( x \in X \). We can then write \( f \leq g + d(f, g) \).

First using monotonicity and then discounting

\[
Tf \leq T\left[ g + d(f, g) \right] \\
Tf \leq Tg + \beta d(f, g) \\
Tf - Tg \leq \beta d(f, g)
\]

By switching \( f \) and \( g \), we analogously find that

\[
-(Tf - Tg) \leq \beta d(f, g)
\]

So we can write that

\[
|Tf_1 - Tg_1| \leq \beta d(f, g) \\
|Tf_2 - Tg_2| \leq \beta d(f, g)
\]

So

\[
\max \{|Tf_1 - Tg_1|, |Tf_2 - Tg_2|\} \leq \beta d(f, g)
\]

Therefore, since this holds for all \( x \in X \).

\[
\sup_{x \in \mathbb{R}^m} \max \{|Tf_1 - Tg_1|, |Tf_2 - Tg_2|\} \leq \beta d(f, g)
\]

This shows that \( T \) is a contraction.

\[
d(Tf, Tg) \leq \beta d(f, g)
\]

---

**Lemma 2.** For any bank, large or small, \( \lim_{x_i \to \infty} L_i(x_i) = \left\{ 1 - p + p\gamma_i \right\} \frac{L_i}{K_i} - \frac{K_i}{p\kappa_i} \) and \( \lim_{x_i \to -\infty} L_i(x_i) = \left\{ 1 - p + p\delta_i \right\} \frac{L_i}{K_i} - \frac{K_i}{p\kappa_i} \).
Proof. Let me restate the optimal response for both large and small banks $L_i(x_i) = \{1 - p + p(\gamma_i + (\delta_i - \gamma_i) \left[ E_i[q(\theta, L)] + L_i \frac{\partial E_i[q(\theta, L)]}{\partial L_i} \right]) \} \frac{r_i^L}{\rho r_i} - \frac{r_i^C}{\rho r_i}$. By noting that, when $x_i \to \infty$, $E_i[q(.)] \to 0$ and when $x_i \to -\infty$, $E_i[q(.)] \to 1$ but again $\frac{\partial E_i[q(.)]}{\partial L_i} \to 0$, it is easy to check that $\lim_{x_i \to \infty} L_i^*(x_i) = \{1 - p + p\gamma_i\} \frac{r_i^L}{\rho r_i} - \frac{r_i^C}{\rho r_i}$ and $\lim_{x_i \to -\infty} L_i^*(x_i) = \{1 - p + p\delta_i\} \frac{r_i^L}{\rho r_i} - \frac{r_i^C}{\rho r_i}$. 

Lemma 3. $\int_{-\infty}^{\infty} \Phi \left( \frac{\theta - x_k}{\sigma_k} \right) d\Phi \left( \frac{x_k - \theta^*}{\sigma_k} \right) = \frac{1}{2}$

Proof. Let’s define $y = \Phi \left( \frac{\theta - x_k}{\sigma_k} \right)$, and $dy = d\Phi \left( \frac{\theta - x_k}{\sigma_k} \right)$. Therefore $\int_{-\infty}^{\infty} \Phi \left( \frac{\theta - x_k}{\sigma_k} \right) d\Phi \left( \frac{x_k - \theta^*}{\sigma_k} \right) = \int_0^1 y dy = \frac{y^2}{2} \bigg|_0^1 = \frac{1}{2}$, where I have used that fact that $\phi (x) = \phi (-x)$.

Lemma 4. $\int \Phi \left( \sqrt{\alpha_y + \alpha_k} \left( \theta^* - \frac{\alpha_y}{\alpha_y + \alpha_k} y - \frac{\alpha_k}{\alpha_y + \alpha_k} x_k \right) \right) \sqrt{\alpha_k} \phi \left( \sqrt{\alpha_k} (x_k - \theta^*) \right) dx_k = \Phi \left( \frac{\alpha_y (\theta^* - y)}{\sqrt{\alpha_y + 2\alpha_k}} \right)$

Proof. It can be shown that $\int_{-\infty}^{\infty} \Phi(A - Bx)\phi(Cx - D)dx = \frac{1}{\sqrt{\pi}} \Phi \left( \frac{C A - B D}{\sqrt{C^2 + B^2}} \right)$. Applying this result:

$$F(\theta) = \Phi \left( \frac{\sqrt{\alpha_k} \cdot \sqrt{\alpha_y + \alpha_k} \theta^* - \alpha_k \sqrt{\alpha_k} \sqrt{\alpha_y + \alpha_k} y - \frac{\alpha_k}{\alpha_y + \alpha_k} \sqrt{\alpha_k} \theta^*}{\sqrt{\alpha_k + \frac{\alpha_k}{\alpha_y + \alpha_k}}} \right) $$

$$= \Phi \left( \frac{\alpha_k \sqrt{\alpha_y + \alpha_k} \theta^* - \alpha_k \sqrt{\alpha_y + \alpha_k} y}{\sqrt{\alpha_k + \frac{\alpha_k}{\alpha_y + \alpha_k}}} \right) $$

$$= \Phi \left( \frac{\alpha_k \sqrt{\alpha_y + \alpha_k} \theta^* - \alpha_k \sqrt{\alpha_y + \alpha_k} y}{\sqrt{\alpha_k (\alpha_y + \alpha_k) + \frac{\alpha_k}{\alpha_y + \alpha_k}}} \right) $$

$$= \Phi \left( \frac{\alpha_k \theta^* - \alpha_k y}{\sqrt{\alpha_y + 2\alpha_k}} \right)$$

7 A discrete action global game formulation

In this section, I briefly describe a version of the model with a binary decision variable in the tradition of Morris and Shin (1998, 2000). From a technical viewpoint, this formulation maps one-to-one to the model analyzed by Corsetti et al. (2004). As in their paper, I only discuss the case with a single large bank $N = 1$.

We have a unit measure of banks, with a measure $1 - \lambda$ corresponding to small banks and a measure $\lambda$ corresponding to a single large bank. Their only choice is to be systemic (high leverage) or nonsystemic (low leverage). The information structure is identical to the one discussed in the main part of the text. The payoffs for the central bank are also identical.
The payoffs for the banks are depicted in figure (1).

Formally, the profit for a bank $i$, where $i = k, j$ is given by

$$
\pi_i = a_i [(1 - p) \pi_S + p \cdot 1 \{B = 1\} \delta \pi_S] + (1 - a_i) \pi_{NS}
$$

(7)

where $a_i = 1$ denotes the action of becoming systemic and $B = 1$ the action of bailing out the banks.

A key condition in this formulation is the following

$$
(1 - p) \pi_S + p \delta \pi_S \geq \pi_{NS} \geq (1 - p) \pi_S \geq 0
$$

(8)

This assumption induces a pecking order in the optimal decisions for the banks: if they know that they are going to be bailed out for sure, they prefer to systemic ($(1 - p) \pi_S + p \delta \pi_S \geq \pi_{NS}$); if they know that they are not going to be bailed out for sure, they would choose to be nonsystemic ($\pi_{NS} \geq (1 - p) \pi_S$).

Under common knowledge about $\theta$, there is a unique equilibrium when $\theta \leq \lambda$ and $\theta > 1$. In the region $\theta \in [\lambda, 1]$ there is multiplicity. Under imperfect information, when $\lambda = 0$ the model is a standard global game: the region of multiplicity disappears and each small bank follows a threshold strategy. When $\lambda = 1$, we have a single agent decision problem and there is also an optimal threshold strategy.

The most interesting is the case with $\lambda \in (0, 1)$. Corsetti et al. (2004) show that there is a unique equilibrium in thresholds strategies for small banks and for the large bank. Their analytical results are derived in the limiting case with $\sigma_x \rightarrow 0$ and $\sigma_y \rightarrow 0$, and $\frac{\sigma_x}{\sigma_y} \rightarrow r$. Under these conditions, it can be shown that the bailout region is large when the large agent is in the economy. Both the size effect and information effect work in the desired direction. Far from the limit it is not possible to prove results analytically, but the mechanism discussed in the paper still holds.
Note that this formulation encompasses similar strategic concerns, but is harder to map to reality. Overall, my continuous variable formulation allows for a much richer set of predictions than the classic binary action game.

8 Additional specifications and residuals

8.1 Additional figure with $\sigma_j = 1$ and $\sigma_k = 1$

![Probability of bailout for different $\lambda$](image.png)

Figure 2: Specification: $p = 1$, $\sigma_j = \sigma_k = 1$ with $N = 1$

This figure is identical to the left plot in figure 9 in the text, with the exception that $N = 1$ instead of $N = 2$. It illustrates how giving more than proportional weight to large banks in $\tilde{L}$, by setting a large $w_j$, would boost the probability of bailout.

8.2 Specification with $\sigma_j = 1$ and $\sigma_k = 2$

The additional plots that complete the analysis of figure 10 in the paper are presented in figures 3 and 4.
Any irregularities in the left pots of figures 3 and 6 are due to numerical error. Note nonetheless than the difference between equilibrium policies is always positive and larger than the maximum collocation error.

8.3 Specification with $\sigma_j = 2$ and $\sigma_k = 1$

For completeness, I present the specification with better informed small banks, despite its lack of realism. I plot here the analogous plots to figures 5 to 8 in the paper.
Figure 5: Specification: \( p = 1, \lambda = 0.75, \sigma_j = 2, \sigma_k = 1, N = 2 \)

Figure 6: Specification: \( p = 1, \lambda = 0.75, \sigma_j = 2, \sigma_k = 1, N = 2 \)
8.4 Residuals

The collocation residuals for the main calibration in the paper and the two additional ones with asymmetric precisions are shown in figure 8. The fact that they are equioscillating is a desirable feature from a numerical perspective.
Figure 8: Multiple specifications: $\lambda = 0.75$, $p = 1$, $N = 2$, upper left plot $\sigma_j = 1$, $\sigma_k = 1$, upper right plot $\sigma_j = 1$, $\sigma_k = 2$, lower plot $\sigma_j = 2$, $\sigma_k = 1$
References


